

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 12.98 Seconds  
(without alignments)  
718.843 Million cell updates/sec

Title: US-09-995-917A-1

Perfect score: 2027

Sequence: 1 MGNPFIGETISFKPHRSDS.....YVEFKRGMLLEPTKFLD 382

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	655	32.3	472	2	US-08-622-166A-2
2	655	32.3	472	2	US-08-622-166A-4
3	294.5	14.5	492	3	US-08-724-466B-2
4	294.5	14.5	492	4	US-08-882-164D-2
5	292.5	14.4	497	4	US-08-882-164D-32
6	269.5	13.3	497	4	US-08-724-466B-4
7	269.5	13.3	497	4	US-08-882-164D-4
8	178.5	8.8	510	3	US-08-606-505B-66
9	178.5	8.8	510	4	US-09-616-990-66
10	172	8.5	508	4	US-08-991-677-2
11	172	8.5	531	4	US-09-380-420C-2
12	170.5	8.4	495	1	US-08-532-065B-2
13	169.5	8.4	506	1	US-08-313-075A-38
14	169.5	8.4	506	3	US-08-606-505B-65
15	169.5	8.4	506	4	US-09-616-990-65
16	164.5	8.1	496	1	US-08-913-075A-50
17	159	7.8	509	3	US-08-948-564-18
18	159	7.8	521	3	US-08-948-564-14
19	158.5	7.8	490	1	US-08-201-118-7
20	158.5	7.8	490	2	US-08-238-821B-7
21	158.5	7.8	490	2	US-08-238-821B-7
22	156.5	7.7	496	4	PCT-US95-05744-7
23	156	7.7	507	4	US-09-172-339-6
24	156	7.7	507	5	US-08-457-274A-23
25	153	7.5	516	3	PCT-US95-05758-23
26	150	7.4	490	1	US-08-948-564-12
27	150	7.4	490	2	US-08-201-118-11
					Sequence 11, Appl

28	150	7.4	490	5	PCT-US95-05744-11	Sequence 11, Appl
29	149	7.4	490	1	US-08-201-118-5	Sequence 5, Appl
30	149	7.4	490	2	US-08-238-821B-5	Sequence 5, Appl
31	149	7.4	490	5	PCT-US95-05744-5	Sequence 5, Appl
32	147	7.3	507	1	US-08-457-274A-22	Sequence 22, Appl
33	147	7.3	507	5	PCT-US95-05758-22	Sequence 22, Appl
34	146.5	7.2	496	3	US-08-881-784-1	Sequence 1, Appl
35	146.5	7.2	496	4	US-09-292-768-2	Sequence 2, Appl
36	146.5	7.2	496	4	US-09-292-768-64	Sequence 64, Appl
37	146.5	7.2	496	4	US-09-292-768-66	Sequence 66, Appl
38	144.5	7.1	513	3	US-08-948-564-6	Sequence 6, Appl
39	144.5	7.1	520	2	US-09-091-432-2	Sequence 2, Appl
40	143.5	7.1	511	4	US-08-991-677-4	Sequence 4, Appl
41	141.5	7.0	476	1	US-08-313-075A-30	Sequence 30, Appl
42	141	7.0	490	1	US-08-201-118-3	Sequence 3, Appl
43	141	7.0	490	1	US-08-201-118-9	Sequence 9, Appl
44	141	7.0	490	2	US-08-238-821B-3	Sequence 3, Appl
45	141	7.0	490	2	US-08-238-821B-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-08-622-166A-2  
; Sequence 2, Application US/08622166A  
; Patent No. 5952545  
; GENERAL INFORMATION:  
; APPLICANT: KONCZ, CSABA  
; APPLICANT: MATYUR, JAIDEEP  
; APPLICANT: SEZERES, MIKLOS  
; APPLICANT: ALTMANN, THOMAS  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME  
; TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/622,166A  
; FILING DATE: 27-MAR-1996  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SVENSSON, LEONARD R.  
; REGISTRATION NUMBER: 30,330  
; REFERENCE/DOCKET NUMBER: 0147-0153P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 205-8000  
; TELEFAX: (703) 205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ. ID NO. 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 472 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-622-166A-2  
Query Match 32.3%; Score 655; DB 2; Length 472;  
Best Local Similarity 34.5%; Pred. No. 9, 2e-60;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;  
QY 1 MGNPFIGETISFKPHRSDSIGTFLOORVSRGKVKNSICGKAVVSCDGLMFILO 60

```

Db      36 LGPLIETETQILGAVKTEPEPEFIDERVARVYGSVFMTHLGEPTLIFSADETNRFLQN 95
QY      61 EGRKFTSDYPRAMHDIIGKYSLLATGEIHRKLNVIISFINLTKSRPDLHCENSTIS 120
Db      96 EGRKFTSDYPRAMHDIIGKYSLLATGEIHRKLNVIISFINLTKSRPDLHCENSTIS 155
QY      121 ILKSMKRCREVEFEKVEKMFLLSVYNQOLSIKPEDPARLYVLDPLSYMGFSLPIPL 180
Db      156 NLDSWSS--RVLMEAKKTIFFELTVQKMSFDGEMSE-SLRKEYLLVIGFSLPL 212
QY      181 PGCTYNAIKVRSNRNIHQNAII-----EDMNAIREDDELDSIISNED----- 224
Db      213 FSTYRRAIQAR--RKVAEALTVVMKRRBEEEGAEKKDKMLAALLAADGFSDEIYD 270
QY      225 -----EEHAIRAKKGCELLMEDIYOKME 249
Db      271 FLVALLVAGYETSTIMTLAVKFLTEPLPLAOLKEEHEKIRAKMSDSYSLEWSDYRSM 330
QY      250 FTQCVISEALRCGNIVTVHRKATHDIKFEKYYVLPKGMKVFPTAVHLDPSLHENPEEF 309
Db      331 FTQCVVNETLRVANIIGVRRAMTDEIKGYKIPKGMKVFSSRAVHLDPNHKKDARTF 390
QY      310 NPMRWTKT-----TAFGGVRCVCGEGELGKIQIAFFLHHLVLSYRM 350
Db      391 NPMRWQSNVYTTGPSNVFTPEGGGPRCLPGYELARVALSVFLHRLVYGFWSW 441

```

## RESULT 2

```

US-08-622-166A-4
; Sequence 4, Application US/08622166A
; Patent No. 5952545
; GENERAL INFORMATION:
; APPLICANT: KONCZ, CSABA
; APPLICANT: MATHUR, JAIDEEP
; APPLICANT: SZEKERS, MIKLOS
; APPLICANT: ALTMANN, THOMAS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
; TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS I
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,166A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0147-0153P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-622-166A-4

```

Query Match 32.3%; Score 655; DB 2; Length 472;  
 Best Local Similarity 34.5%; Pred. No. 9,2e-60;  
 Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

QY      1 MGWPIGETTSFFKPHNSDSIGTFLOORVSRVYGVKFSNIGRAVYSCDOLMFIION 60
Db      36 LGPLIETETQILGAVKTEPEPEFIDERVARVYGSVFMTHLGEPTLIFSADETNRFLQN 95
QY      61 EGRKFTSDYPRAMHDIIGKYSLLATGEIHRKLNVIISFINLTKSRPDLHCENSTIS 120
Db      96 EGRKFTSDYPRAMHDIIGKYSLLATGEIHRKLNVIISFINLTKSRPDLHCENSTIS 155
QY      121 ILKSMKRCREVEFEKVEKMFLLSVYNQOLSIKPEDPARLYVLDPLSYMGFSLPIPL 180
Db      156 NLDSWSS--RVLMEAKKTIFFELTVQKMSFDGEMSE-SLRKEYLLVIGFSLPL 212
QY      181 PGCTYNAIKVRSNRNIHQNAII-----EDMNAIREDDELDSIISNED----- 224
Db      213 FSTYRRAIQAR--RKVAEALTVVMKRRBEEEGAEKKDKMLAALLAADGFSDEIYD 270
QY      225 -----EEHAIRAKKGCELLMEDIYOKME 249
Db      271 FLVALLVAGYETSTIMTLAVKFLTEPLPLAOLKEEHEKIRAKMSDSYSLEWSDYRSM 330
QY      250 FTQCVISEALRCGNIVTVHRKATHDIKFEKYYVLPKGMKVFPTAVHLDPSLHENPEEF 309
Db      331 FTQCVVNETLRVANIIGVRRAMTDEIKGYKIPKGMKVFSSRAVHLDPNHKKDARTF 390
QY      310 NPMRWTKT-----TAFGGVRCVCGEGELGKIQIAFFLHHLVLSYRM 350
Db      391 NPMRWQSNVYTTGPSNVFTPEGGGPRCLPGYELARVALSVFLHRLVYGFWSW 441

```

## RESULT 3

```

US-08-724-466B-2
; Sequence 2, Application US/08724466B
; Patent No. 6063606
; GENERAL INFORMATION:
; APPLICANT: Petkovich, P. Martin, White, Jay A.,
; APPLICANT: Beckett, Barbara R., Jones, Glenville
; TITLE OF INVENTION: Retinoid Metabolizing Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Blake, Cassels & Graydon
; STREET: Box 25, Commerce Court West
; CITY: Toronto
; COUNTRY: Canada
; ZIP: M5L 1A9
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
; COMPUTER: COMPAQ, IBM PC compatible
; OPERATING SYSTEM: MS-DOS 5.1
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,466B
; FILING DATE: October 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/667,546
; FILING DATE: June 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunt, John C.
; REGISTRATION NUMBER: 36,424
; REFERENCE/DOCKET NUMBER: 50767/00004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 863-4344
; TELEFAX: (416) 863-2653
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-724-466B-2

```

```

Query Match Similarity      14.5%: Score 294.5; DB 3: Length 492;
Best Local Similarity      24.3%: Pred. No. 3.5e-22;
Matches 111; Conservative   59; Mismatches 190; Indels    97; Gaps    13

Oy  1 MGMPFGETISFFKPHRSDSIGTFLOQRRSRGRGVFKNSICGKAVVSCDCELMETFLON 60
Db  49 |||||L-----|||:::||||:|::|::|::|::|::|::|::|::|::|::|
Oy  61 ECKLETSTOVPRKAMHDLCYKSLLATGELIHRKLKNVIYSFINUTSKRPDLCAENLSIS 120
Db  104 EKKLVSVOPASVRFTLGSDTLSNVHGYOKKKRAIMAFS--RDALHYIPVIOOEVS 162
Oy  121 IILSW---KNCREVEHFHKVKMFTLSVMNQLSIKPE---DPARLYVLDFLSYMKGF 173
Db  163 AIGEWLQKNSC--VLYPEMKKLMRIAMRILLGEPEDIKTDEQL--VEAFEEIKNL 218
Oy  174 ISLPPLPETGTGNIAKVAS-----NNIHONAIIEDMNNAIRBEDFLDSITSN---ED 224
Db  219 FELPFDVPFGSGYKGLRANKNFTHSKIIEENIRKKIODDDNENQKYKDALQLLENSRSD 278
Oy  225 EEHAAIRAKKGDEL-----LNMEVDYOK-----247
Db  279 EEFSLQAKKEATTELLFGCHETTASTATSLVWFLGINTLEVOKVARENOEYKEMGMTPRG 338
Oy  248 -----MEFTOCVISSEALRCGNIVKTVRKATHDIKEFYVIPKGMVFPIPTAVH 297
Db  339 KGLSMELLDTQLKYTCGVIKETLRINRPVGGFRVALKTFELNGYDIPKMWNTYSISCDPH 398
Oy  298 LDRSLHENFEENPNRNRYKT-----TAFGGGVRCPCPGSELGKDIDIAFLHHLVLS 347
Db  399 DVADVPFNKEEPORPERFSKGLEDSRFNYIPFGGSRCVCGERPAKVLIKFIVELTQH 458
Oy  348 YRMKIKSDEM-----PIAHPYVEFRGMLLEIEPKF 379
Db  459 CMMILSNCPPTMKKTGPTIY-----VDNLPIKF 486

RESULT      4
US-08-882-164D-2
: Sequence 2, Application US/08882164D
: Patent No. 6306624
: GENERAL INFORMATION:
: APPLICANT: Petkovich, P. Martin, White, Jay A.,
: APPLICANT: Beckett, Barbara R., Jones, Glenville
: TITLE OF INVENTION: Retinoid Metabolizing Protein
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Blake, Casels & Graydon
: STREET: Box 25, Commerce Court West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5L 1A9
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
: OPERATING SYSTEM: MS-DOS 5.1
: SOFTWARE: WORD PERFECT
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/882,164D
: FILING DATE: June 25, 1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/667,546
: FILING DATE: June 21, 1996
: APPLICATION NUMBER: 08/724,466
: FILING DATE: October 1, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Hunt, John C.
: REGISTRATION NUMBER: 36,424
: REFERENCE/DOCKET NUMBER: 50767//00010
: TELECOMMUNICATION INFORMATION:
```

APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
APPLICATION NUMBER: 08/724,466  
FILING DATE: October 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-882-164D-32

Query Match 14.4%; Score 292.5; DB 4; Length 497;  
Best Local Similarity 23.2%; Pred. No. 5,8e-22;  
Matches 106; Conservative 61; Mismatches 187; Indels 103; Gaps 10;

QY 1 MGPFIGETISFPRHSDSIGTFLQORVSRYGKFKSNICGKAVVSCDQELNMFILN 60  
DB 49 MGPFPEETLQWVLRK-----FLQMKRRKRGPIYKTHLFGRTVAVMGADVNRILLG 103  
QY 61 EGRKFTSDYKRAMHDLIGKYSLLATGEIHRKLNVIISFINLTKSPDFLHCENLSTIS 120  
DB 104 EHRIVSVWMPASVYTLIGACLSNLDHSDSHKORRKYIMQAFSREALOCYVLAEEVS-S 162  
QY 121 ILSMKNC--REVEFKKEMFTLSVMNQLSIKP-----EDPARLYVLODFLSYMKG 172  
DB 163 CLEWMLSCGERGLLYVEVKRMFRIMRILLCGEPGAGGEDQOL--VEAFEEETR 220  
QY 173 FLSPLPFGTYNAIKVSNRNHONAIIDMNNAIIE-----EDFLDSIISN 222  
DB 221 LFLSLPIVPSGLYGVKA--RNLHARIEENIRAKIRLQATEPDGCKDALOLLIEH 277  
QY 223 -----EDEHHAIRA 232  
DB 278 SWGERGLDMQALKOSTELLFGHETTASANSLLITGLYPLVLOKRVREILSKGLC 337  
QY 233 KRGGDELNMEDYOKMEFTOCVISEALRCGNIVKVRKATHDIFKEEYVPRKGMKVEPI 292  
DB 338 KSNQDNKLDMEITLQKLYICVIKETLRNPPVPGFRVALKTFELNGYQIPKGMNVIYS 397  
QY 293 FTAVHLDPSSLHENPFEFNPWKMT-----KTAFGGGVAVCPGGELGKQIAFFLH 342  
DB 398 ICDTHDVADIFTNKEEFNPDRFIVPHEDASRSFSIFPGGLRSCVGEFAKILKIFTV 457  
QY 343 HLVLSTYRWKIKS-----DEMPLAHRYVE 365  
DB 458 ELARHCDWQLNGPPTMKTSPTVYPVNDLPARFTYFQ 494

RESULT 6  
US-08-724-466B-4  
Sequence 4, Application US/08724466B  
Patent No. 6063606  
GENERAL INFORMATION:  
APPLICANT: Petkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
ZIP: M5L 1A9  
COUNTRY: Canada  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAO, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,466B  
FILING DATE: October 1, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-466B-4

Query Match 13.3%; Score 269.5; DB 3; Length 497;  
Best Local Similarity 23.2%; Pred. No. 1.4e-19;  
Matches 100; Conservative 64; Mismatches 176; Indels 91; Gaps 13;

QY 1 MGPFIGETISFPRHSDSIGTFLQORVSRYGKFKSNICGKAVVSCDQELNMFILN 60  
DB 49 MGPFPEETLQWVLRK-----FLQMKRRKRGPIYKTHLFGRTVAVMGADVNRILLG 103  
QY 61 EGRKFTSDYKRAMHDLIGKYSLLATGEIHRKLNVIISFINLTKSPDFLHC-----AE 115  
DB 104 DDLVSVWMPASVYTLIGSGLSNLDHSDSHKORRKYIMRFS-----REALCYVPIYE 158  
QY 116 NLSISILSKWNC--REVEFKKEMFTLSVMNQLSIKP-----EDPARLYVLODFLSY 169  
DB 104 DDLVSVWMPASVYTLIGSGLSNLDHSDSHKORRKYIMRFS-----REALCYVPIYE 158  
QY 159 EVG-SLLEQWMLSCGERGLLYVEVKRMFRIMRILLCGEPQLAGDSDSQIVAEFEEM 217  
DB 170 MKGFTSLPFGTYNAIKVSNRNHONAIIDMNNAIIE-----EDFLDSIISN 222  
QY 218 TRNLSPLPIVPSGLYGVKA--RNLHARIEENIRAKIRLQATEPDGCKDALOLLIEH 277  
QY 205 -----DMNNAIIE-----DEFLDSIISN-----NEDEHHAIRA 231  
DB 278 SWGERGLDMQALKOSTELLFGHETTASANSLLITGLYPLVLOKRVREILSKGLC 336  
QY 232 AKKGDGELNMEDYOKMEFTOCVISEALRCGNIVKVRKATHDIFKEEYVPRKGMKVEPI 291  
DB 337 KSNQDNKLDMEITLQKLYICVIKETLRNPPVPGFRVALKTFELNGYQIPKGMNVIYS 396  
QY 292 FTAVHLDPSSLHENPFEFNPWKMT-----KTAFGGGVAVCPGGELGKQIAFFLH 341  
DB 397 SICDTHDVADIFTNKEEFNPDRFIVPHEDASRSFSIFPGGLRSCVGEFAKILKIFT 456  
QY 342 HLVLSTYRWKI 352  
DB 457 VELARHCDWQL 467

RESULT 7  
US-08-882-164D-4  
Sequence 4, Application US/08882164D  
Patent No. 6306624  
GENERAL INFORMATION:  
APPLICANT: Petkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:



ADDRESS: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5L 1A9  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,164D  
FILING DATE: June 25, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
APPLICATION NUMBER: 08/724,466  
FILING DATE: October 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-882-164D-4

Query Match 13.3%; Score 269.5; DB 4; Length 497;  
Best Local Similarity 23.2%; Pred. No. 1.4e-19;  
Matches 100; Conservative 64; Mismatches 176; Indels 91; Gaps 13;

QY 1 MGNPFGTISFPKPHRSISIGTFLOQVRSRYGKVFKNICGKAVSCDDDELNFILON 60  
DB 49 MGNPFGTISFPKPHRSISIGTFLOQVRSRYGKVFKNICGKAVSCDDDELNFILON 103  
QY 61 ECKLFTSDVPKAMHDLIGKYSLLATGELHRLKNVYISFINLTKSKPDLHC-----AE 115  
DB 104 DDLVSVHMPASVKTILSGCISNLHDSHKOKKVIKMAFS-----REALECYVYITE 158  
QY 116 NLISILSKSMKNC--REVEFHKEVKMFTLSVNVNQLSIRP---EDPARLYVLODFLSY 169  
DB 159 EYV-SSLEQMLSCGERGLLVYEPVKRLMFRIMARILLGCEPOLAGDGDSEQLVEAFEEEM 217  
QY 170 MKCGFISLPPLPGCTGTNAIKVAS--NRNIHON-----AIIIE- 204  
DB 218 TRNLFSLPIDVPSGLRGKARNLHARLEQIRAKICGLRASAGGCKDALQLLIEH 277  
QY 205 -----DMNNAIRREE-----DFLDISIIS-----NEDEHAIR 231  
DB 278 SWRGERLDM-QALKOSSTELLFGHETASATSLITVLGLYPHVLOKVRLELKKGLL 336  
QY 232 AKKGDELLNMEDEYQKMEFTQCVISBALRCGNIVKTVHRKATHDIFKEVYIPKGVKVP 291  
DB 337 CKSNODKIDMELDELOKYGIVKIEFLRLNRPVPGGFVALKTFELNGYQIPKGNVYI 396  
QY 292 IFTAVHLDSLHNEPPEFNMRMT-----KTTAFGGCVRVCPGGGLGLOLAFL 341  
DB 397 SICDTHDVAIEFTNKEFEFNDRESAPAPEDASRFSFLPEGGGLRSCGKEFAKILKIF 456  
QY 342 HHLVLSYRWKI 352  
DB 457 VELARHCMDQL 467

RESULT 8

US-08-606-505B-66  
Sequence 66, Application US/08606505B  
Patent No. 6114601  
GENERAL INFORMATION:  
APPLICANT: KIKUCHI, Yasuhiro  
APPLICANT: KIKUCHI, Yasuhiro  
APPLICANT: KIKUCHI, Yasuhiro  
APPLICANT: SHIMADA, Yukihisa  
APPLICANT: OHBAYASHI, Masaya  
APPLICANT: SHIMADA, Ritsuko  
APPLICANT: OKINAKA, Yasushi  
TITLE OF INVENTION: NOVEL PLANT GENES  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112-3801  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Ver3.30  
SOFTWARE: PATENT AID Ver1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,505B  
FILING DATE: 23-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP44965/92  
FILING DATE: 02-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Perry, Lawrence S.  
REGISTRATION NUMBER: 31865  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-218-2100  
TELEFAX: 212-218-2200  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 510 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Eustoma russellianum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 92 to 1621  
IDENTIFICATION METHOD: by experiment  
US-08-606-505B-66

Query Match 8.8%; Score 178.5; DB 3; Length 510;  
Best Local Similarity 20.5%; Pred. No. 4.5e-10;  
Matches 95; Conservative 68; Mismatches 160; Indels 141; Gaps 21;

QY 1 MGNPFGTISFPKPHRSISIGTFLOQVRSRYGKVFKNICGKAVSCDDDELNFILON 57  
DB 42 ICMPIVGLARLGLTMH-----VALANNAKKYGPVMTLKVQSC-GLAVASIPBAKAR- 93  
QY 58 LONEGLFTSDVPK--AMHDLIGKYSLLA---TGEHRLKLNVIISINLTKSKPDL 111  
DB 94 LKTDLMNFSNRPNGATHLAVNAQDMVPADYCPKRLKLSNIHI----- 140  
QY 112 HCAENLISILSKSMKCREVEF-----H-----KEVKFTLSVNVNQLSIR 153  
DB 141 -----LGKALGQWEVERKKEGLYMAAESGRHGOVYVSDMLTYAMANNLGOVMLSK 195  
QY 154 PEDPARLYVLO-----DSYMKGFTSLPPLPGCTGTNAIKVASN-RNIHQ-- 199  
DB 196 -----RVFGSGGSESNFEDMNVELMTVAGYFNIGOFIPISIAMMDLOGIGMKRLHKKF 250  
QY 200 -----NAIIEDMNNAIRREE-----DFLDISIISNEDEHAIA-----IKKD 236

Db 251 DALLFRLEEHHTASAEHRKSPDFLDFVYVANGDNGSBERLQTVNITAKALLNMTAGTDTIS 310  
QY 237 GELLNW-----EDYQKMEFTQCVISEALR-CGNI 264  
Db 311 SSVIEWALAEELKNPILIRRAOEMDGVIGRDRFLEADISKLPLYQAIKCEAFKRHPST 370  
QY 265 VKTVHRKATHDIKFEKVEYIPKGMKVPPIFTAVHLDPSSLHNEPFEENPMKWT----- 315  
Db 371 PLNLPRIASQACEVNGHYIPKGTSLSVNITWAIGRDPSVWENPNEPNDPRLERKNAKIDP 430  
QY 316 -----KTAFFGGGVRCVCPGEGELGKLOIAFELHNLVLSYRWKIKS 354  
Db 431 RGNDFELIPFGAGRICAGTRIGILLVEYITLGTIVHSFWELPS 474

RESULT 9  
US-09-616-990-66  
; Sequence 66, Application US/09616990  
; Patent No. 6232109  
; GENERAL INFORMATION:  
; APPLICANT: KIKUCHI, Yasuhiro  
; KIKOKAWA, Shigeto  
; SHIMADA, Yukihisa  
; OHBAVASHI, Masaya  
; SHIMADA, Ritsuko  
; OKINAKA, Yasushi  
; TITLE OF INVENTION: NOVEL PLANT GENES  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112-3801  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.  
; COMPUTER: IBM PS/4  
; OPERATING SYSTEM: MS-DOS Ver3.30  
; SOFTWARE: PATENT AID Ver1.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/616,990  
; FILING DATE: 14-Jul-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP44963/92  
; FILING DATE: 02-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perry, Lawrence S.  
; REGISTRATION NUMBER: 31865  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-218-2100  
; TELEFAX: 212-218-2200  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 510 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Eustoma russellianum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 92 to 1621  
; IDENTIFICATION METHOD: by experiment  
; SEQUENCE DESCRIPTION: SEQ ID NO: 66  
US-09-616-990-66

Query Match 8.8%; Score 178.5; DB 4; Length 510;  
Best Local Similarity 20.5%; Pred. No. 4.5e-10;  
Matches 95; Conservative 68; Mismatches 160; Indels 141; Gaps 21;

QY 1 MGPPIFGE-TISFEKPHRSDSIGTEFLQOVRYSRYKV--EKSNIQCGKAAVSCDQELNMEI 57  
Db 42 IGPVVLGALRLIGTMRP-----VALANNAKKYGPWYLVKQSC-GLAVASPEAKAP- 93  
QY 58 LONEGKLFITSDPK--AMHDIIGKYSLLA--TGEIHRKLKNVYIISINLTKSPDPL 111  
Db 94 LKTLDMNFSNRPNACATHLATNAODMVFADYGPWKLLRKLKLSNHI----- 140  
QY 112 HCAENLISISILKSWKNCREVER-----H-----KEVKFTLSVAVNOLLSTIK 153  
Db 141 -----LGRKALQGWEEVRRKKEGLGYMLYAAESGRHQPPVYSEMLTYANANLGGVMLSK 195  
QY 154 PEDPARLVYQ-----DF-----LSYMGFTSLPIPLDGTGTNAYIKRSN-RNIHQ-- 199  
Db 196 -----RVFGSQSESNEFDNMYELMTVAGYFNIGDFTPSIAMDLOGIGGMKRLHKKF 250  
QY 200 -----NAIIEDMNNATREE-----DFLDSITSNDEEHA-----IRA-----KKGD 236  
Db 251 DALLFRLEEHHTASAEHRKSPDFLDFVYVANGDNGSBERLQTVNITAKALLNMTAGTDTIS 310  
QY 237 GELLNW-----EDYQKMEFTQCVISEALR-CGNI 264  
Db 311 SSVIEWALAEELKNPILIRRAOEMDGVIGRDRFLEADISKLPLYQAIKCEAFKRHPST 370  
QY 265 VKTVHRKATHDIKFEKVEYIPKGMKVPPIFTAVHLDPSSLHNEPFEENPMKWT----- 315  
Db 371 PLNLPRIASQACEVNGHYIPKGTSLSVNITWAIGRDPSVWENPNEPNDPRLERKNAKIDP 430  
QY 316 -----KTAFFGGGVRCVCPGEGELGKLOIAFELHNLVLSYRWKIKS 354  
Db 431 RGNDFELIPFGAGRICAGTRIGILLVEYITLGTIVHSFWELPS 474

RESULT 10  
US-08-991-677-2  
; Sequence 2, Application US/08991677A  
; Patent No. 6252135  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Vincent L  
; APPLICANT: Cartaway, Daniel T  
; APPLICANT: Smeltzer, Richard H  
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms  
; FILE REFERENCE: 50617  
; CURRENT APPLICATION NUMBER: US/08/991,677A  
; EARLIER APPLICATION NUMBER: US 60/033,381  
; EARLIER FILING DATE: 1997-12-16  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Liquidambar styraciflua  
US-08-991-677-2

Query Match 8.5%; Score 172; DB 4; Length 508;  
Best Local Similarity 25.4%; Pred. No. 2.1e-09;  
Matches 52; Conservative 31; Mismatches 60; Indels 62; Gaps 7;

QY 186 TNALIKR-----SNRNHONALIEDMNNATIREDFLDSITSNDEEHAIRAKKGDE 238  
Db 303 TTAISVENAMAEELIKNRVQOKA-----QEELDNVLGSE-----R 337  
QY 239 LNMWEDYQKMEFTQCVISEALRCGNIYKTVH-----RKATHDIKFEKVEYIPKGMKVF 290  
Db 338 VLTLEDFSSLPYLCVAKEAR-----LHPPTPLM.PHRANANVRIKGYDIPKGSNVH 390  
QY 291 PIFTAVHLDPSSLHNEPFEENPMKWT-----KTAFFGGGVRCVCPGEGELGKLOIAF 339  
Db 391 VNVAVARDPAVWBDPLEFFPERFSEDDVDMKGDHYRLLPFGAGRVCVPGAQOLGINLVTS 450

OY 340 ELHVLVSYRWK---IKSDMPIA 360  
DB 451 MMGHLHHFYWSPKGVKPEIDMS 475

## RESULT 11

US-09-380-420C-2  
Sequence 2, Application US/09380420C  
Patent No. 6300544  
GENERAL INFORMATION:  
APPLICANT: Hakler, Barbara  
Bak, Soren  
Kahn, Rachel  
Moller, Birger  
TITLE OF INVENTION: Cytochrome P450 Monooxygenases  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Syngenta Patent Dept.  
CITY: RTP  
STREET: 3054 Cornwallis Road  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/380,420C  
FILING DATE: 12-NO. 6300544-1999  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-21251A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 531 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-380-420C-2

Query Match 8.5%; Score 172; DB 4; Length 531;  
Best Local Similarity 22.3%; Pred. No. 2.3e-09;

Matches 87; Conservative 65; Mismatches 142; Indels 96; Gaps 17;

OY 50 DQELMFLIQ-----NECKLFTSDYPKAMHD-ILGKYSLLATGEIH-----RLK 94  
DB 172 EQEMRLVADLDRAAKASIVLNDHFALTDGIIIG-----YVAFGNITASQOFARKENQ 227  
OY 95 NVIISFINL--TKSRPDLHCAENISILSKWKNCREVEFHKEVKMFTLSVYNQLLSI 152  
DB 228 HVLDDAMOMMSFSAEDFPNNAAGRLADRLSGFLARRERIFN-ELDVFEEKYIDQHM--- 283  
OY 153 KREDPARLYVLDPLSYMKGFISLPIPLPGCYTNAI-----KYRSNNITQNNII 203  
DB 284 ---DAR-----FVPDNGDVLVDVLTINLCKEHDGTLRFTRD-HVKAIV 322  
OY 204 ED-----MNNAREEDFLDSIISNEDEHAAIRAKKGDGE-LNWMEDYQ 246  
DB 333 LDTFGAIDTSSVTILMMAMSELMRKPYLV-----RKAQEVRAVAVDDKPRVNSEDA 375  
OY 247 KMEFTQCVISEALRC-GNIVTVHKKATHDIKREYVPRKMKVPRPITVAHLDPSSLHEN 305  
DB 376 KIPYKAVYKELRLHAPATLLVPREYMRDITIGYDVPANTRVFVNMAAIGRDPASMPA 435  
OY 306 PFEFNPMMT-----KTFAFGGVAVRCPGELGKLIQIAFFLHHLVLSYRWKI-- 352

DB 436 PDEFNDRFVSGDVYGYGSHFELIPFGACRRICPGLTNGETVNTFTLANLTCYDMALPG 495  
OY 353 ---KSDMPIAH-PVEFKRGMLEIEPTKF 379  
DB 496 AMKPEDVSMEETGALTFRHKRTPLVVVPTRY 525

## RESULT 12

US-08-532-065B-2  
Sequence 2, Application US/08532065B  
Patent No. 5753507  
GENERAL INFORMATION:  
APPLICANT: Ohta, Daisaku  
Mizutani, Masaharu  
TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5753507artis Corporation  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: NJ  
COUNTRY: USA  
ZIP: 07936  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/532,065B  
FILING DATE: 22-SEP-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 495 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-532-065B-2

Query Match 8.4%; Score 170.5; DB 1; Length 495;  
Best Local Similarity 19.7%; Pred. No. 2.9e-09;  
Matches 88; Conservative 65; Mismatches 180; Indels 113; Gaps 14;

OY 4 PFIETISFFR-PIRSDSIGTFLQORVSRGKVKFSNIGGKAVASCDQELMFTIQLNEG 62  
DB 46 PIGNIHVLGKHPRHS-----FAELSKTYGPVMSILGSLNTVIVASPEARAVLRTHD 99  
OY 63 KLFTSDVP-----KAMHDLGKYSL-----LATGSIHKKL 94  
DB 100 QILASRPTNAVRSININODASLVLPSSASARWILRLRSTYQLISPORIETKRLRNKV 159  
OY 95 NVIISFINLTKSRPDLHCAENISILSKWKNCREVEFHKEVKMFTLSVYN--QLLSI 152  
DB 160 KELVSFISESDRESVDISRVAFTITLNTISN---ILFSVDGSYNAKKSINGVODTVI 216  
OY 153 KREDPARLYVLDPLSYMKGFISLPIPLPGCYTNAIKYSNNRIH-----QNAITE 204  
DB 217 SYMDAGCPDPAANFPLRL-FLDL-----QGNVTKETVCTERLVRVPRGTDAKIAEKS 269  
OY 205 DMNNA--IREEDFLDSI-----ISNEDEH----- 227  
DB 270 SQNNPKDVSXNDPVDNLDLYKGDSELSISDIEHLLDMFTAGTDISSITLWPMTELLK 329

QY 228 -----AARAKKGBELINMEDYOKMEFTQCVISEALRCGNIVK-TVHRRKATHDIX 277  
DB 330 NPKTAKAOAEIDCVIGONGIVEESDISKPLQAVKETFRLHPVPLLIPRKESDAE 389  
QY 278 FKEYIIPKGVKVFPIFTAVHLDPSLHNPFEFNPWRMT-----KTYAFGGVAV 326  
DB 390 ILGFVNLDTQVLVNVMAIGRDPVWDNPSQEPERFLGKDMYGRDVELPFGAGRI 449  
QY 327 CPGBELGKLIQIAFLHHLVLSYRWKI 352  
DB 450 CPGMPLAMKTVSIMLASLISYDFMRL 475  
RESULT 13  
US-08-313-075A-38  
Sequence 38, Application US/08313075A  
Patent No. 5639870  
GENERAL INFORMATION:  
APPLICANT: Holton, Timothy A.  
APPLICANT: Cornish, Edwin C.  
APPLICANT: Tanaka, Yoshikazu  
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID  
TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/313,075A  
FILING DATE: 30-NOV-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PL 1538/92  
FILING DATE: 27-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PL 6698/93  
FILING DATE: 07-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PCT/AU93/00127  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9433  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 506 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-313-075A-38  
Query Match 8.4%; Score 169.5; DB 1; Length 506;  
Best Local Similarity 21.3%; Pred. No. 3.9e-09;  
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;  
QY 2 GMPFIGE-TISFKRPHRSDSIGFLQORVSRKGV--FKSNIGCGRAVAVSCDELINLFI 58  
DB 39 GMPVIGALPILGAMPVHS-----LAKMAKKYGAIMYLKVGTC-GMAVASTPDAKAF-L 90

QY 59 ONEGKLTSDYPK--AMEDIIGKYSLLA-----TGEIHRKLKNVIISFINLTKSPDELI 112  
DB 91 KTLIDFNFSNRPENAGATLHATNAQDMFAHYGPRKLLRLKSN-----LH 135  
QY 113 CAENLSISILSKWKNCREVER-H-----REVKMTLSVNVNOLLSTKP 154  
DB 136 M---LGKALEMNAVVRANLGLHMLKSMDMSREGQRYVAEMLETFANANNIGQVLMKSK- 191  
QY 155 EDPARLIY-----LQDF-----LSYMGFISLPPLPGTGTN--AIKVSNNRIHQ-- 199  
DB 192 ---RVFVDKGYEVNEFKDMVVELMTIAGYFNIGDIFCLAMMDQGLEKRRKR-LHKKF 246  
QY 200 NAIIDMNNAIR-----EEDFDSIISNEDEEHA-----IRA----- 232  
DB 247 DALLTKMEDEKATTYERKGRKDFLDVYMENDNSEGRSLSTTNKALLNLFTAGTDS 306  
QY 233 -----KKGDEL-----LNMEDYOKMEFTQCVISEALRCG--- 262  
DB 307 SSAIEMWALAEMKKNPAILKKAQEMDQVIGRRNRILLESIDIPNLPLYRAIKETEFKHPST 366  
QY 263 --NIVKTVHRRKATHDIXKREYIIPKGVKVFPIFTAVHLDPSLHNPFEFNPWRMT----- 315  
DB 367 PLNLPRISNEPCIVD---GYIIPKNTRLSVNIMAIGRDPQVWENPLERPERFLSGNS 422  
QY 316 -----KTYAFGGVAVCPGBELGKLIQIAFLHHLVLSYRWKIKSD 355  
DB 423 KIDPGRNDELIPFGAGRICTAGTGMIVMEYIIIGTLVHSDFMRLPSE 471  
RESULT 14  
US-08-606-505B-65  
Sequence 65, Application US/08606505B  
Patent No. 6114601  
GENERAL INFORMATION:  
APPLICANT: KIRUCHI, Yasuhiro  
APPLICANT: KIROKAWA, Shigetaro  
APPLICANT: SHIMADA, Yukihisa  
APPLICANT: OHBAYASHI, Masaya  
APPLICANT: SHIMADA, Ritsuko  
APPLICANT: OKINAKA, Yasushi  
TITLE OF INVENTION: NOVEL PLANT GENES  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELIA, HARPER & SCINTO  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112-3801  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.50 inch, 720 KB storage.  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Ver3.30  
SOFTWARE: PATENT AID Ver1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,505B  
FILING DATE: 23-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP44963/92  
FILING DATE: 02-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Perry, Lawrence S.  
REGISTRATION NUMBER: 31865  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-218-2100  
TELEFAX: 212-218-2200  
INFORMATION FOR SEQ ID NO: 65 :  
SEQUENCE CHARACTERISTICS:  
LENGTH: 506 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Petunia hybrida  
STRAIN: Falcon Blue  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 116 to 1633  
IDENTIFICATION METHOD: by experiment  
US-08-606-5058-65

Query Match  
Best Local Similarity 8.4%; Score 169.5; DB 3; Length 506;  
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;

QY 2 GMPFGE-TISFFKPHRSDSIGTFLQORVSRGKV-FKSNICGKAVVSCDDELNMFIL 58  
DB 39 GMPVIGALPILGAMPNHS-----LAKMAKKYGAIMLKVGTC-GMAVASTPDAKAF-L 90  
QY 59 QNEKLTSDYPR--AMHDILGKYSLLA---TGEIHRKLKNVIISFINTLSKDPDLCH 112  
DB 91 KTLIDINSNRPNGATHLAVNADWVFAYHGRMKLLRLSN-----LH 135  
QY 113 CAENLISILSKMKNCREVEF-H-----KEYKMTLSVMVQULSIRP 154  
DB 136 M---LGKALENNANVRANLGHILKMSDMSREGORVVAEMLTETAMNMIGOVMLSK- 191  
QY 155 EDPARLVY-----LQDF-----LSYMKGFISLPPLPGTYN--AIKVRNRIHQ-- 199  
DB 192 ---RFEVDKGVEVNEFKDMVVELMTIAGYFNIGDFIPCLAMMDLOGIEKRMKR-LHKKF 246  
QY 200 NAIIEDNNNAIR-----EEDFLDSIISNEDEHNA-----IRA----- 232  
DB 247 DALLTKMFDEHKATTYERKKRPDLVYVMEGNSGGERLSTTNIKALLNLTACTDTS 306  
QY 233 -----KKGDEL-----LNMEDYQKMEFTQCVISALRCG--- 262  
DB 307 SSAIEWALAEKMKNPAILKKAQAEKMDQVIGRNRLLESIDIPNLPYRAICKETFRKHPST 366  
QY 263 --NIVTVHRKATHDIKFEYVPRKGMKVPFTAVHLDPSLHENFEENPRKMT----- 315  
DB 367 PLNLPRIISNEPCIVD---GYIIPKNTRLSVNIMALGRDPOVWENPLENPERFLSGRNS 422  
QY 316 -----KTTAFGGGVAVCPGEGELGKQLIAFFLHNLVLSYRWIKSD 355  
DB 423 KIDPRGNDFELIPFGAGRICAGTRMGIVMVEYIIGTLVHSDWMLPSE 471

RESULT 15  
US-09-616-990-65  
Sequence 65, Application US/09616990  
Patent No. 6232109

GENERAL INFORMATION:  
APPLICANT: KIKUCHI, Yasuhiro  
KITOKAWA, Shigetaro  
SHIMADA, Yukihisa  
OHBAAYASHI, Masaya  
SHIMADA, Ritsuko  
OKINAKA, Yasushi  
TITLE OF INVENTION: NOVEL PLANT GENES  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112-3801

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.  
COMPUTER: IBM PS/4  
OPERATING SYSTEM: MS-DOS Ver3.30  
SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/616,990  
FILING DATE: 14-Jul-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP44963/92  
FILING DATE: 02-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Perry, Lawrence S.  
REGISTRATION NUMBER: 31865  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-218-2100  
TELEFAX: 212-218-2200  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 506 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Petunia hybrida  
STRAIN: Falcon Blue  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 116 to 1633  
IDENTIFICATION METHOD: by experiment  
SEQUENCE DESCRIPTION: SEQ ID NO: 65  
US-09-616-990-65

Query Match  
Best Local Similarity 8.4%; Score 169.5; DB 4; Length 506;  
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;

QY 2 GMPFGE-TISFFKPHRSDSIGTFLQORVSRGKV-FKSNICGKAVVSCDDELNMFIL 58  
DB 39 GMPVIGALPILGAMPNHS-----LAKMAKKYGAIMLKVGTC-GMAVASTPDAKAF-L 90  
QY 59 QNEKLTSDYPR--AMHDILGKYSLLA---TGEIHRKLKNVIISFINTLSKDPDLCH 112  
DB 91 KTLIDINSNRPNGATHLAVNADWVFAYHGRMKLLRLSN-----LH 135  
QY 113 CAENLISILSKMKNCREVEF-H-----KEYKMTLSVMVQULSIRP 154  
DB 136 M---LGKALENNANVRANLGHILKMSDMSREGORVVAEMLTETAMNMIGOVMLSK- 191  
QY 155 EDPARLVY-----LQDF-----LSYMKGFISLPPLPGTYN--AIKVRNRIHQ-- 199  
DB 192 ---RFEVDKGVEVNEFKDMVVELMTIAGYFNIGDFIPCLAMMDLOGIEKRMKR-LHKKF 246  
QY 200 NAIIEDNNNAIR-----EEDFLDSIISNEDEHNA-----IRA----- 232  
DB 247 DALLTKMFDEHKATTYERKKRPDLVYVMEGNSGGERLSTTNIKALLNLTACTDTS 306  
QY 233 -----KKGDEL-----LNMEDYQKMEFTQCVISALRCG--- 262  
DB 307 SSAIEWALAEKMKNPAILKKAQAEKMDQVIGRNRLLESIDIPNLPYRAICKETFRKHPST 366  
QY 263 --NIVTVHRKATHDIKFEYVPRKGMKVPFTAVHLDPSLHENFEENPRKMT----- 315  
DB 367 PLNLPRIISNEPCIVD---GYIIPKNTRLSVNIMALGRDPOVWENPLENPERFLSGRNS 422  
QY 316 -----KTTAFGGGVAVCPGEGELGKQLIAFFLHNLVLSYRWIKSD 355  
DB 423 KIDPRGNDFELIPFGAGRICAGTRMGIVMVEYIIGTLVHSDWMLPSE 471

Search completed: July 29, 2002, 13:58:24  
Job time: 51 sec

---

## ALIGNMENTS

Tue Jul 30 08:40:07 2002

us-09.

RESULT 1  
ID 09LY89 PRELIMINARY: PRT: 382 AA.  
AC 09LY89;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 43.9 KDA PROTEIN.  
GN F18022.190.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,  
RA Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL: AL163817; CAB87779.1; -;  
DR InterPro: IPR001128; Cyt\_P450.  
DR Pfam: PF00067; P450; 1.  
DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Hypothetical protein; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 382 AA: 43889 MW: 185655AB73B4E30 CRC64;

Query Match 99.6%; Score 2018; DB 10; Length 382;  
Best Local Similarity 99.5%; Pred. No. 1,1e-153;  
Matches 380: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCMPIGIGTISFFKPHRSOSIGTFLQORVSRVRYGKFKSNICGKAVVSCDQELNMFTLQN 60  
DB 1 MCMPIGIGTISFFKPHRSOSIGTFLQORVSRVRYGKFKSNICGKAVVSCDQELNMFTLQN 60

QY 61 EKLFTSDYPKAMHDILGKYSLLATGELIRKILKNVIISPIINTKSKPDLCAENLSTIS 120  
DB 61 EKLFTSDYPKAMHDILGKYSLLATGELIRKILKNVIISPIINTKSKPDLCAENLSTIS 120  
QY 121 ILKSMKNCREVEFEFKKMTLSVYNQULSTIKPEDPARLYVLODFLSYMKGFISLPPL 180  
DB 121 ILKSMKNCREVEFEFKKMTLSVYNQULSTIKPEDPARLYVLODFLSYMKGFISLPPL 180  
QY 181 PGTGYTNAIVRSNRNIHONAIIEDMNNAIREDPFLDSTISNDEDEHAAIRAKKGDGELL 240  
DB 181 PGTGYTNAIVRSNRNIHONAIIEDMNNAIREDPFLDSTISNDEDEHAAIRAKKGDGELL 240  
QY 241 NMEDYOKMEFTQCVISEALRCGNIYVYRKATHDIKREYVYIPGKRVFPIETAVHNDP 300  
DB 241 NMEDYOKMEFTQCVISEALRCGNIYVYRKATHDIKREYVYIPGKRVFPIETAVHNDP 300  
QY 301 SLHENPEEPNPMRWTKTTFAGGVRVCPGSELGKLOIAFLLHLVLSYRWKIKSDENPIA 360  
DB 301 SLHENPEEPNPMRWTKTTFAGGVRVCPGSELGKLOIAFLLHLVLSYRWKIKSDENPIA 360  
QY 361 HPYVEFKRGMLEIEPTKFLD 382  
DB 361 HPYVEFKRGMLEIEPTKFLD 382

- Support yield this is it aligned?  
- Support function based on spec  
- % 500 is up function

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 : Search time 28.42 seconds  
(without alignments)  
2325.266 Million cell updates/sec

Title: US-09-995-917a-1  
Perfect score: 2027  
Sequence: 1 MGFPIGETISFEKPHRSDS.....YVEFKRGMLEIETPFLIED 382

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mnc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP\_rvirus:\*  
16: SP\_bacteriap:\*  
17: SP\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2018	99.6	382	10	09LY89
2	822	40.6	513	10	064989
3	821	40.5	513	10	09SC09
4	642.5	31.7	474	10	09LKH7
5	599.5	29.6	512	10	09FX29
6	592.5	29.2	491	10	0941A6
7	586	28.9	490	10	0941W5
8	573	28.3	464	10	0941C5
9	572	28.2	465	10	0940V4
10	572	28.2	478	10	09LW73
11	569	28.1	465	10	09LW73
12	546	26.9	465	10	09FM45
13	513.5	25.3	467	10	0949P1
14	513	25.3	463	10	09FI38
15	505.5	24.9	463	10	09FM76
16	504	24.9	457	10	065624

17	487.5	24.1	485	10	09S0H2	09sjh2 arabidopsis
18	486	24.0	443	10	09LJK2	09ljrk2 arabidopsis
19	484	23.9	482	10	081077	081077 arabidopsis
20	468.5	23.1	735	10	09IG17	091g17 arabidopsis
21	467	23.0	486	10	09LIG9	091lig9 arabidopsis
22	460	22.7	455	10	09LXH8	091lxh8 arabidopsis
23	431.5	21.3	477	10	09LVY7	091vy7 arabidopsis
24	427	21.1	489	10	09ZV72	09zlv72 arabidopsis
25	427	21.1	489	10	09CSY2	09csy2 arabidopsis
26	423	20.9	490	10	09CSY3	09csy3 arabidopsis
27	422.5	20.8	464	10	004949	004949 arabidopsis
28	422.5	20.8	474	10	09SHY7	09shy7 arabidopsis
29	417.5	20.6	496	10	09FGY4	09fgy4 cucurbita m
30	416	20.5	487	10	023384	023384 arabidopsis
31	413	20.4	497	10	09AXW6	09axw6 taxus cuspi
32	412.5	20.4	460	10	09SVX2	09svx2 arabidopsis
33	408.5	20.2	504	10	09SNC3	09sng3 oryza sativ
34	376	18.5	499	10	09AXH9	09axh9 hordeum vul
35	292.5	14.4	497	11	09R1F4	09r1f4 mus musculu
36	283.5	14.0	224	10	09SDM6	09sdm6 helianthus
37	280.5	13.8	375	10	09LN32	09ln32 arabidopsis
38	280	13.8	492	13	093323	093323 xenopus lae
39	278.5	13.7	512	4	09NR63	09nr63 homo sapien
40	276	13.6	525	4	09NP41	09np41 homo sapien
41	275	13.6	492	13	09PUB4	09pub4 gallus gall
42	224.5	11.1	494	11	09QZ50	09qz50 mesocricetu
43	223.5	11.0	494	11	091XG2	091xg2 mus musculu
44	217.5	10.7	494	11	091X75	091x75 mus musculu
45	215.5	10.6	349	10	09FMR7	09fmr7 arabidopsis

#### ALIGNMENTS

RESULT 1  
09LY89 ID 09LY89 PRELIMINARY: PRT: 382 AA.  
AC 09LY89;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOHETICAL 43.9 KDA PROTEIN.  
GN F18022\_190.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,  
RA Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL, AL163817; CAB87779.1; -  
DR InterPro, IPR001128; Cyt\_P450.  
DR Pfam, PF00067; p450; 1.  
DR PROSITE, PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme: Hypothetical protein; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 382 AA; 43889 MW; 1B65685AB73BAE30 CRC64;

Query Match 99.6%; Score 2018; DB 10; Length 382;  
Best Local Similarity 99.5%; Pred. No. 1.1e-153;  
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGFPIGETISFEKPHRSDSIGTFLQORVSRGKVFKNICGKAVVSCDGLNMFILION 60  
DB 1 MGFPIGETISFEKPHRSDSISTFLQORVSRGKVFKNICGKAVVSCDGLNMFILION 60

```

QY 61 ECKLETSOYPRKAMHDLGKYSLLANGELTHRKLNKNTSPINLTGSKPDLHCENLSTIS 120
DB 61 ECKLETSOYPRKAMHDLGKYSLLANGELTHRKLNKNTSPINLTGSKPDLHCENLSTIS 120
QY 121 ILSKMKNCREVEFEHKEVKTSLVYNQLLSIKPEDPARLYLQDPLSYMKGFSIPDL 180
DB 121 ILSKMKNCREVEFEHKEVKTSLVYNQLLSIKPEDPARLYLQDPLSYMKGFSIPDL 180
QY 181 PGTGTNAIKVSNRNHONAIIEEDMNNAIREDPDLSTISNEDEHAAIRAKKGDELL 240
DB 181 PGTGTNAIKVSNRNHONAIIEEDMNNAIREDPDLSTISNEDEHAAIRAKKGDELL 240
QY 241 NMEDQOKMEFTQCVISEALRCGNIVYTHRKATHDIKFEKVEYIPKGMKVPITFAVHLP 300
DB 241 NMEDQOKMEFTQCVISEALRCGNIVYTHRKATHDIKFEKVEYIPKGMKVPITFAVHLP 300
QY 301 SLHENPFEENPMRWKTTFAGGGVRCVPCGELGKLIQIAFFLHHLVLSYMKIKSDMP1A 360
DB 301 SLHENPFEENPMRWKTTFAGGGVRCVPCGELGKLIQIAFFLHHLVLSYMKIKSDMP1A 360
QY 361 HPYEFKRGMLLEIETKLEED 382
DB 361 HPYEFKRGMLLEIETKLEED 382

```

# RESULT 2

```

ID 064989 PRELIMINARY; PRT; 513 AA.
AC 064989;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE.
GN DWF4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WS-2;
RX MEDLINE=98158690; PubMed=9490746;
RA Choe S., Dilkes B.P., Fujioka S., Takatsuto S., Sakurai A.,
RA Feldmann K.A.;
RT "The DWF4 gene of Arabidopsis encodes a cytochrome P450 that mediates
RT multiple 22alpha-hydroxylation steps in brassinosteroid
RT biosynthesis";
RL Plant Cell 10:231-243(1998).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF044216; AAC05093.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA; 58867 MW; B1639BDD9A4DA6F3 CRC64;

```

Query Match 40.6%; Score 822; DB 10; Length 513;  
 Best Local Similarity 34.4%; Pred. No. 1, 1e-57;  
 Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;

```

QY 2 GMPFGETISFPRKPHRSDSITGFLOQVRVRYKGVKFSNCGKAVVSCQOELMFLQNE 61
DB 45 GMPFGETISFPRKPHRSDSITGFLOQVRVRYKGVKFSNCGKAVVSCQOELMFLQNE 61
QY 62 GKLFTSDYPRKAMHDLGKYSLLANGELTHRKLNKNTSPINLTGSKPDLHCENLSTIS 121
DB 105 GKLFTSDYPRKAMHDLGKYSLLANGELTHRKLNKNTSPINLTGSKPDLHCENLSTIS 121
QY 122 ILSKMKNCREVEFEHKEVKTSLVYNQLLSIKPEDPARLYLQDPLSYMKGFSIPDL 181
DB 122 ILSKMKNCREVEFEHKEVKTSLVYNQLLSIKPEDPARLYLQDPLSYMKGFSIPDL 181

```

```

DB 165 LDSMOONSIFSAQDEAKFTFNLMAKHMSMDGEEBETOLKKEYVTFMKGVVADPLNP 224
QY 182 GTGTNAIKVSNRNHONAIIEEDMNNAIREDPDLSTISNEDEHAAIRAKKGDELL 240
DB 225 GTATYHAKLOSRAITILKFEIERKMERKLDIKKEEQEEBEVKTDEAEKMSDHYKQRTDD 284
QY 225 -----EEHAAI 230
DB 285 DLGLWYKHSNSTEQILDLISLFAGHETSSVAIALAIFLQACPKAVEELREHLEI 344
QY 231 -RAKKGDELL-LNEDYQKMEFTQCVISEALRCGNIVYTHRKATHDIKFEKVEYIPKGMK 288
DB 345 ARAKKELGESLELWDDYQKMEFTQCVISEALRCGNIVYTHRKATHDIKFEKVEYIPKGMK 404
QY 289 VPIPTFAVHLPDLHENPFEENPMRWKTTFAGGGVRCVPCGELGKLIQIAFFLHHLVLSYMK 329
DB 405 VLPVISAVALHDSRYDQPLNFPWRMOQOONGASSSGSGSFYWGNNYMPFGGGPRLCAG 464
QY 330 GELGKLIQIAFFLHHLVLSYMKIKSDMP1AHPYEFKRGMLLEI 374
DB 465 SELAKLEMAVFIHHLVLSYMKIKSDMP1AHPYEFKRGMLLEI 509

```

# RESULT 3

```

ID 09SC09 PRELIMINARY; PRT; 513 AA.
AC 09SC09;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE (DWF4) (AT3G50660/73A5_40).
GN 73A5_40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bloeker H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortum M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davits R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AL132979; CAB62435.1; -.
DR EMBL; AF412114; AAL06567.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00365; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA; 58867 MW; B1639BDD9A5D7C93 CRC64;

```

Query Match 40.5%; Score 821; DB 10; Length 513;  
 Best Local Similarity 34.4%; Pred. No. 1, 3e-57;  
 Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;

```

OY 2 GMPFGETISFKPHRSISIGTFLQORVSRGKVKFSKNICGGKAVVSCDQELNMFLLONE 61
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 45 GMPFGETISGKPYATATLGDPMOQHVSKYKYSRNLFGPPTIYSADAGLNRLFIONE 104
OY 62 GKLFTSDYKAMHDLGKSLLATGEIRKLKNVYISINLTGKSPDLHCAENLISLI 121
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 105 GKLFTSDYKAMHDLGKSLLATGEIRKLKNVYISINLTGKSPDLHCAENLISLI 164
OY 122 LKSMKNCREVEFHKEVKMFTLSVMYNOLISIKREDPARLYVLDLSYMKGFISLPPLP 181
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 165 LDMQONSTFSSADDEKKTFTFLMAKHIMSDRGEETQLKKEYVTGPKGVASAPLNP 224
OY 182 GCGYNAIKVRSNRNIHONAIEDNNAIREDLDSITSNED----- 224
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 225 GTAHKAALOSRATILKETERKMEERKLDKEDEEEVEKTEDEAEMSKDHVRKORTD 284
OY 225 -----DEHAI 230
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 265 DLGAVLKHNSITQIIDLILSLFAGHEISSVAIALAIFLQACPKAVEELREHLEI 344
OY 221 -RAKKGDEL-LNMBDYOKMEFTQCVISALRCGNIVTVHRKATHDIKFEYVIPKGMK 288
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 345 ARAKKEGESESLNMDYKMKDETCVINTLRLGNVRFRLHKKALKDVRKGYDIPSGMK 404
OY 289 VRFITAVHLDPSLHENPPEFPMRWTKT-----AFGGVRYVCPG 329
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 405 VLPVISAVALHDSRYDQPLFNPMPMOONNGASSGSGSEFTMGNNYMPFGGPRLAG 464
OY 330 GELGKLQIAFLHLLVLSYRMKIKSDMPIAHYVEFKGMLEI 374
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 465 SELAKLEMAVFIHLLVLFKNMELADDPKFAFPVDFPGLIRV 509

```

```

RESULT 4
O9LKH7 PRELIMINARY: PRT: 474 AA.
AC O9LKH7:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME P450.
GN CYP90A2.
OS Vigna radiata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_TaxID=15791;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. 2937.
RA Yang M.T., Chen Y.M.;
RT Cloning and sequencing of a Vigna radiata cDNA encoding cytochrome
RT P450.
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AF279252; AAF89209.1;
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 2.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 474 AA; 54037 MW; A2542A809C5BACD CRC64;

```

```

Query Match 31.7%; Score 642.5; DB 10; Length 474;
Best Local Similarity 32.7%; Pred. No. 2,4e-43;
Matches 140; Conservative 71; Mismatches 140; Indels 77; Gaps 6;

```

```

Db 97 GKLDCSVPGSISNLGKHSLLMKGALKRKHSLTMSANSIIRKDLHHLIDRLIGLN 156
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
OY 122 LKSMKNCREVEFHKEVKMFTLSVMYNOLISIKREDPARLYVLDLSYMKGFISLPPLP 181
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 157 LDTWSD--RVTLMDQAKITFELTVKOLMSFDPDEWTE-SLRKEVVLVIEGFFLLPLF 213
OY 182 GCGYNAIKVRSNRNIHONAIEDNNAIREDLDSITSNED-----EFDLSITSNED----- 224
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 214 STTYRAIKARK-----VAEALLTVVRQREYNOGKEKKSMLALLAGSGHFSDD 266
OY 225 -----EEHAIIRAKKGDELNMBDY 245
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 267 QLVDFLLALVAGYETSTIMTLAVKFLTEPRLAQLKEHNDQIRASDCAPLEMTDY 326
OY 246 QKMEFTQCVISALRCGNIVTVHRKATHDIKFEYVIPKGMKVRPITAVHLDPSLHEN 305
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 327 KSWFTQHVNETLKVANIIGGIFRRATTDIDIKGTYTPKGMKVFASPRAVHLNPEYYKD 386
OY 306 PFEFPMRWTKT-----TAFGGVRYVCPGSELGKLQIAFLHLLVLSYRMKIKS 354
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 387 ARTFPMRMOSSSRAANPANYTTFGGGPRLCFGEIARVYLVFLHRIYTRSWPAE 446
OY 355 DEMPIAHP 362
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 447 EDKLVFFP 454

```

```

RESULT 5
O9FX29 PRELIMINARY: PRT: 512 AA.
AC O9FX29:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE, PUTATIVE.
GN T9L24.44.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Renning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.U., White O., Nierman W.C., Fraser C.M.;
RT Arabidopsis thaliana chromosome 1 BAC T9L24 genomic sequence.
RC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC EMBL: AC012396; AAC30983.1;
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 512 AA; 58297 MW; 1590C0304BF319FA CRC64;

```

```

Query Match 29.6%; Score 599.5; DB 10; Length 512;
Best Local Similarity 30.6%; Pred. No. 7,4e-40;
Matches 144; Conservative 74; Mismatches 139; Indels 113; Gaps 15;

```

```

OY 2 GMPFGETISFKPHRSISIGTFLQORVS-----RYGVK 36
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 49 GMPLIGDFPALNVAAGSHSSFEVKQIKFVSLSCVLLILKRPNSGPNETIRYGRIF 108
OY 37 KSNICGKAVVSCDQELNMFLLONEGKLTSDYKAMHDLGKSLLATGEIRKLKNV 96
    ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 109 SCSLFGKMAVVASADPDRNRFIMONEGKLFQSSYKSRDLVGKGVITVHGDQORLHSI 168
OY 97 IISFINLTGKSPDLHCAENLISILKSMKN-----CREVEFHKEVKMFTLSVMVN 147

```

```
Db 169 ASSMRRHDLKTHFLVPIPVNMLQTLISNFKDEEVLIDODICKVAIH-----LAWN 219
Qy 148 QLLSIKPEDPARLYLQDFLSYMKGFISLPILPGTYNAIKVNS-----NRNIH--- 198
Db 220 QALLGSSSEBVD-EMSQLEFSDVDGLSVIPIDLPGETYKAKAKKEIRKINKTIEKRL 278
Qy 199 QALLIED-----MNNAREE-----DFLDSIT--SNED----- 224
Db 279 QKKAASDFTAGNGLRLGLEESLIPNESMADFTINLFAENETTSKTMLEPAVYELTHCPKA 338
Qy 225 -----EEHAIRAKKGGDELLNMEYOKMEFTQCVISEALRCGNIVKTVHRKATHDIKFK 279
Db 339 MTQLLEEHRLA-----GGMLTMDTKTMDFTQCVIDEFTLRIGLAIWLMREAKEDEVSTQ 393
Qy 280 EYVIPKGMKVPFLPTAVHLDPSLHENPEFNPMRW-----KTT-----AEGGVR 325
Db 394 DYVIPKGCVFVFLSAVHLDESRYKESLSFNPMRWLDPETQOKRMWRTSPFCFGGTR 453
Qy 326 VCPGGLGLQIAFLHLVLYSTRW-KTSDPMPIAHPIYVEKRGMLLEI 374
Db 454 FCPGALARIQIALFLHFIITYKWTQLEKEDRISF-FPSARLVNGFKIOL 502
```

```
RESULT 6
Q94IAG PRELIMINARY; PRT; 491 AA.
AC Q94IAG;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CYP90D.
GN CYP90D.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RA Shimada Y.;
RT Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB066286; BAB62109.1; - F450 gene repressed by the brassinosteroid.*;
SQ SEQUENCE 491 AA; 56153 MW; 02PB908A91995A40 CRC64;
```

Query Match 29.2%; Score 592.5; DB 10; Length 491;  
Best local Similarity 31.4%; Pred. No. 2.6e-39;  
Matches 137; Conservative 75; Mismatches 160; Indels 65; Gaps 8;

```
Qy 1 MGMPFGETISFEKPHRSDSIGTFLQORVSRYGKVKNSICGKAVYSCDQELNFTLON 60
Db 56 LGMPVIGETIEFVSASPRPESFMDKRLMYGRVFSHFFGTATVSDAEVNRVLQS 115
Qy 61 EGGLEFTSDPKAMHDILGKYSLLATGEIRKLVNIIISINLTGSKPDLHCAENLSIS 120
Db 116 DSTAFVPEFTKYRELKMGSSILLINGSILHRRHGLVGSFLKSPLLKAQIVRDMHFLSE 175
Qy 121 ILKSMKNCREVEFHKEKMFETLSVMVNOQLSI-KPEDPARLYLQDFLSYMKGFISLP 179
Db 176 SMQLESDQGVLLQDYSKTVAFKVLAKALISVEKGEDEL--KREPEFISGLMSLPIN 233
Qy 180 LPETGTNAL-----KVSNRNIIHONAIT-----EDMNNAI 210
Db 234 FPGTQHRSLQAKRMVQYERIIEGKIRTKKKEEDVDYAKVDVLLKDSSEHLTHNL 293
Qy 211 REEDFDLSIISNED-----EEHAIRA-KKGGDELNMEY 245
Db 294 IANNMIDMLIPGHDSPVVLITLAVKFLSDPALNLTENMKLSLAKELTGPRLWYNDY 353
Qy 246 OKMEFTQCVISEALRCGNIVKTVHRKATHDIKREYVIPKGMKVPFLPTAVHLDPSLHEN 305
```

```
Db 354 ISLPTQKVTITETLMKGNVITIGMKRAKDVIRKGVIPKGCFLIATLRSVHLDTLYES 413
Qy 306 PEPFPMRW-----TKTAEGGVAVVCPGSELGKLQIAFLHLVLSYRKIKSDMP 358
Db 414 PYKFNPMRWQERDMNTSSPFGGQRLCPGIDLARLETSTVSLHLVTRFW-IAEDDTI 472
Qy 359 IAHPIYVEKRGMLLEI 375
Db 473 INFPTVHMKNKPLWIK 489
```

```
RESULT 7
Q94IWS PRELIMINARY; PRT; 490 AA.
AC Q94IWS;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CYTOCHROME P450-LIKE PROTEIN.
GN P0419B01.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0419B01."
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003244; BAB56089.1; - EC2853BBAFB88F CRC64;
SQ SEQUENCE 490 AA; 54824 MW; 9EC2853BBAFB88F CRC64;
```

Query Match 28.9%; Score 586; DB 10; Length 490;  
Best local Similarity 30.4%; Pred. No. 8.5e-39;  
Matches 131; Conservative 87; Mismatches 153; Indels 60; Gaps 9;

```
Qy 2 GMPFGETISFEKPHRSDSIGTFLQORVSRYGK-VKNSICGKAVYSCDQELNFTLON 60
Db 55 GMPVIGETIEFVSASPRPESFMDKRLMYGRVFSHFFGTATVSDAEVNRVLQS 114
Qy 61 EGGLEFTSDPKAMHDILGKYSLLATGEIRKLVNIIISINLTGSKPDLHCAENLSIS 120
Db 115 DARAFVPEFTKYRELKMGSSILLINGSILHRRHGLVGSFLKSPLLKAQIVRDMHFLSE 174
Qy 121 ILKSMKNCREVEFHKEKMFETLSVMVNOQLSIKPEDPARLYLQDFLSYMKGFISLP 180
Db 175 ALSSFPDSSLHVQHLAKSVFEILVRGLGLEAGEEMQ-QLKQCPQEFIVGIMSLPIKL 233
Qy 181 PGTYNAIKVNS-----NRNIH-----NIIEDMNNAREE---DF 215
Db 234 PGTRLYRSLOAKKMARLQRLIRERARRAASPPRAIDVLIGSDSELDELISDMN 293
Qy 216 IDSIIISNED-----EEHAIRAKKGD-CELLNMEYOKMEF 250
Db 294 IDLMIPEDSVPVLLITLAVKFLSECPDLAHLQLEENIQOKRRTDMDGELTQWMDVSLSF 353
Qy 251 TQCVISEALRCGNIVKTVHRKATHDIKREYVIPKGMKVPFLPTAVHLDPSLHENPEFN 310
Db 354 TQHVITETLRGNIIIGIMKKAADVGVNHLIPKGMCVFVFRSVHLDLDTLYDEPYKRN 413
Qy 311 PMRWTKT-----TAFGGVAVVCPGSELGKLQIAFLHLVLSYRKIKSDMP 363
Db 414 PMRWKEDMSNGSFPTFGGQRLCPGIDLARLETSTVSLHLVTRFW-IAEDDTI 472
Qy 364 VEFKRGMLLEI 374
Db 473 VRLKRGPIRV 483
```

```

RESULT 8
09LNC5 PRELIMINARY: PRT: 464 AA.
AC 09LNC5
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME P450-LIKE PROTEIN.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA:
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA:
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety p1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AP001307; BAB01922.1;
DR InterPro: IPR001126; CYL_P450.
DR InterPro: IPR000504; RRM.
DR Pfam: PF00067; P450.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 464 AA; 52864 MW; A65E09465E5190B CRC64;

Query Match 28.3%; Score 573; DB 10; Length 464;
Best Local Similarity 31.6%; Pred. No. 8.7e-38;
Matches 130; Conservative 70; Mismatches 147; Indels 64; Gaps 7;

QY 1 MGWPIGTFISFFKPHRSDSICGFLQQRVSRGKFKFNICGKAVNSCDDQNLNFIION 60
DB 56 LGMPVIGETIEFVSAYSDDRESFMDKRLMYGKFKSHIGTATIVSTDAEVNRAVLQS 115
QY 61 ECKLFTSDYPRKAMHDLCKYSLLATGEIHRKLVNITISFINLTKSKDPFLHCAENLSIS 120
DB 116 DSTAFVPEPPTVRLMKSSILLNGSLHRRHGLVGSFLKSPLLKAQIVADMHKFLSE 175
QY 121 ILKSKNCRVEYEFHKEVFMFLSVVNNQLST-KPEDPARLYVLQDFLSYMGFISLPIR 179
DB 176 SMDLMSDDQPLVDQVSTVAFKYLAKALISYEKEDLEEL-KREFEFISGLSLPLIN 233
QY 180 LPTGQYTNAI-----KVRSNRNINQNAII-----EDMNNAI 210
DB 224 FPGQGLHNSLQAKKNMVKOVERIEIGKIRKTKNKEEDVIAKDVVLLKSSSEHLTNL 293
QY 211 REEDFLDSITISNED-----BEHAATA-KKGGDELLNMEDEY 245
DB 294 IANNIMDMIGHGSVPVLLTAVLFKSDSPAALLNLTRENNKTKSLKELGCEPLVMNDY 353
QY 246 OKMEFTQCVISEALRCGNIVTVHKKATHDIKFKEVYIPKGMKVPPIPTAVHLDPSLHN 305
DB 354 LSLPTQAVITETLLMGVNIIGVMKAKMDVEIKGYVIPKGCPLAYIRSVHLDKLYES 413
QY 306 PEEFNPMK-----TKTAFGGGVGVVCPGGELGKQIAEFLHNLVLSYR 349
DB 414 FYKFPNPMQERDMNTSSFSFGGQGRICPGDLARLETSTVFLHNLVYRFR 464

```

```

RESULT 9
0940V4 PRELIMINARY: PRT: 465 AA.
AC 0940V4
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE AT3G30180/T20F20.6.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA:
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamuya A., Katlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortolmi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY052655; AAK96559.1;
SQ SEQUENCE 465 AA; 53814 MW; 280A21D0712FCAL1 CRC64;

```

```

Query Match 28.2%; Score 572; DB 10; Length 465;
Best Local Similarity 30.7%; Pred. No. 1e-37;
Matches 136; Conservative 71; Mismatches 152; Indels 84; Gaps 9;

```

```

QY 1 MGWPIGTFISFFKPHRSDSICGFLQQRVSRGKFKFNICGKAVNSCDDQNLNFIION 60
DB 39 MGWPIGTFETTERLKQGPD-----FMKNQRLRYGSFFKSHIICPTIVSDAELNRIILMN 93
QY 61 ECKLFTSDYPRKAMHDLCKYSLLATGEIHRKLVNITISFINLTKSKDPFLHCAENLSIS 120
DB 94 ESKGLVAGVPOSMIDLIGCNIAVHGSPSHRLMRGSLSLISPTMKDKLLPKRIDPFMN 153
QY 121 ILKSKNCRVEYEFHKEVFMFLSVVNNQLST-----KPE-DPARLYVLQDFLSYMKCFI 174
DB 154 YLGGWDDLETVDIOEKTHM--AFLSILQIAETLKRPREVEYR---TEFFKLVVGL 206
QY 175 SLPIPLPGYGVNATKVRNINQNAIIEDMNNAIREE-----DELDSITISNED---- 224
DB 207 SVPIDIPGNTNRSGVQARNINIDRLITELMQE-----RKSGETPMGLGKMKEDNRYL 261
QY 225 -----EBHAAIRAKGDELLN 241
DB 262 LTKDEIRQDVVITLGSYETVSTSMALKYLHDPKALELRHRLAIRKKRRDEPLT 321
QY 242 WEDYQKMEFTQCVISEALRCGNIVTVHKKATHDIKFKEVYIPKGMKVPPIPTAVHLDPS 301
DB 322 LDDIKSMKFTVAFETSLATVINGVLRKTHNDELNGYLIRPKGMIRVYVREINQYTS 381
QY 302 LHENFEFNPMKWTKT-----AFGGGVGVVCPGGELGKQIAEFLHNLVLSRWKTK 353
DB 362 LTEDPMIFNPMKWMKKSLSKSYFLFLFGGVRLCPGKELGISEVSSFLHYFTYKRWEN 441
QY 354 SDEMPIAHYVEFFKRGMLEIEP 376
DB 442 GEDKLMVFPRVSAPKGYHLKCSF 464

RESULT 10
09LNT3 PRELIMINARY: PRT: 478 AA.
AC 09LNT3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```



DE CYTOCHROME P450 (BRASSINOSTEROID-6-OXIDASE).  
 GN BR60X1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxId=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=98290546; PubMed=9628582;  
 RA Sato S., Kaneo T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,  
 RA Tabata S.,  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.  
 RT Physically assigned P1 and TAC clones.";  
 RT DNA Res. 5:41-54(1998).  
 RL (2)  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=21295570; PubMed=11402205;  
 RA Shimada Y., Fujioke S., Miyauchi N., Kushiro M., Takatsuto S.,  
 RA Nomura T., Yokota T., Kamiya Y., Bishop G., Yoshida S.,  
 RT "Brassinosteroid-6-oxidase from Arabidopsis and tomato catalyze  
 RT multiple C-6 oxidations in brassinosteroid biosynthesis.";  
 RL Plant Physiol. 126:770-779(2001).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL: AB009048; BAB08653.1; -;  
 DR EMBL: AB005868; BAB08658.1; -;  
 DR InterPro: IPR001128; CYL\_P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 KM  
 SQ SEQUENCE 465 AA; 53767 MW; 0C00459C9C866D1F CRC64;

Query Match 26.98; Score 546; DB 10; Length 465;  
 Best Local Similarity 29.78; Pred. NO. 1.3e-33;  
 Matches 130; Conservative 73; Mismatches 161; Indels 74; Gaps 9;

1 MGMPFGETISFFKPHRSDSIGTFLQORVSRGKVFKNICGKAVYSCDQELMFLION 60  
 39 MGMPFGETISFFKPHRSDSIGTFLQORVSRGKVFKNICGKAVYSCDQELMFLION 93  
 61 ESKLETSYPRAMHDLICKYSLLATGEIHRKLNVIISFTNLTKSKPDLHCAENLSIS 120  
 94 ESKGLVPYQSMIDLICTCMMAVHGSSHRLMGSLSLISSTMHDLIPKVDHFMRS 153  
 121 ILKSKNCREVEFEKVEKMFSLVNVQDLS--IKPEDRALYVQ---DPLSTMKGFIS 175  
 154 YLDQNNLEVEDIDQDKTHNAFLSSLTQIAGNLKRP-----FVEERKTAFLKLVGTL 207  
 176 LPIPLPGTGYNAIKVRSNRNIHONAIIEDMNA-----IREEDF 215  
 208 VPILPGTNYRCGIQARNRINRDLRELWQRODSGETFDMLGKMKEGRIYLTBEI 267  
 216 LDIIS-----NEDEHAATAKKGDGELLNWDYOK 247  
 268 RDQVVTILYSGEYVSTSMALKYLHDHPKALQELRAEHLAFERRKRODEPLLEDYKS 327  
 248 MEPTQVISEALRCGNIVKTVHRKATHDIKFEKVIIPKGMKVPPIFAVHLDPSLHNP 307  
 328 MKFRAVIYERSRLATVNGVLRKTRDELNGILPKGRITYYTRITNDANLYEDPL 387  
 308 EENMRWTKT-----AFGGVAVCPGEGELKQIAFLHLHLVLSYRW-KIKSDMP 358  
 388 IFNFMRMKKSLESQNSCFVGGGTGLCPGKELGIYESSFLHYFYRYRMEEGDEL- 446  
 359 IAHYVEFKRGMLEIEP 376  
 447 MVPRVAVAPKGFHRTISP 464

RESULT 13  
 ID 0949P1 PRELIMINARY; PRT; 467 AA.  
 AC 0949P1;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PUTATIVE CYTOCHROME P450 PROTEIN.  
 GN ATG19230.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxId=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,  
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,  
 RA Tang C., Tortum M., Yamamura Y., Yu G., Yu S., Bowser L.,  
 RA Carinici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,  
 RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shion P., Southwick A., Tracy S.E.,  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,  
 RT "Full length cDNA of gene Atg19230 (GI:7268718)."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY050980; AAK93657.1; -;  
 SQ SEQUENCE 467 AA; 53037 MW; 2F4230446536D955 CRC64;

Query Match 25.38; Score 513.5; DB 10; Length 467;  
 Best Local Similarity 29.48; Pred. No. 5.1e-33;  
 Matches 126; Conservative 69; Mismatches 150; Indels 83; Gaps 13;

1 MGMPFGETISFFKPHRSDSIGTFLQORVSRGKVFKNICGKAVYSCDQELMFLION 60  
 41 MGMPFGETISFFKPHRSDSIGTFLQORVSRGKVFKNICGKAVYSCDQELMFLION 96  
 61 ESKLETSYPRAMHDLICKYSLLATGEIHRKLNVIISFTNLTKSKPDLHCAENLSIS 117  
 97 KSHLEKTFEPASKERMGKQAIFFHOGDYHAKRLKRLAFMPESIRNVPDI-----ESI 152  
 118 SISLSKMKCREVEFEKVEKMFSLVNVQDLSI-----KPEDRALYVQDPLSTY 170  
 153 AODSLRSMWETM--INTQEKTKYTFNV---ALSLTGKDEVLRYEDLKRCYVLE----- 203  
 171 KPTSLDPIPLPGTGYNAIKVRSN-RNIHONAIIEDMNAIREEDFLDIISNED----- 224  
 204 KGYNSMVPVNLPGTLFHKSMKARELSQILARILSERQNGSSHNDLGSFMDGKELTDE 263  
 225 -----BEHAATAKKGDGELLNWDY 245  
 264 QIADNIICTVFAARDTASVSMILKYLAENPVNLEAVTEBOAIIKDEEGSLTWGDT 323  
 246 QKMEFTQVISEALRCGNIVKTVHRKATHDIKFEKVIIPKGMKVPPIFAVHLDPSLHNP 305  
 324 KKMPLTSRYIQTFLRAVASTLSFTFREAVDEVEGTLIPKGMKVPPIFAVHLDPSLHNP 383  
 306 PFEFNMRW-----TKTAFGGVAVCPGEGELKQIAFLHLHLVLSYRWK--KSDM 357  
 384 PGKFDPSREVAAPKPTNPFNGHSHSCPNELAKLEMSIMIHNLTKYSMSIVASDGI 443  
 358 ---PIAHP 362  
 444 QYGFALP 451

RESULT 14  
 ID 09FI38 PRELIMINARY; PRT; 518 AA.  
 AC 09FI38;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)

01-MAR-2001 (Tremblrel. 16, last sequence update)  
 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE CYTOCHROME P450-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 Miyajima N., Tabata S.,  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT pl and TAC clones."  
 RL DNA Res. 6:183-195(1999).  
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AB017064; BAB1064.1;  
 DR InterPro; IPR001128; Cyt\_P450.  
 DR Pfam; PF00067; P450; 2.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 518 AA; 58877 MW; CCA3F74F4301563E CRC64;

Query Match 25.3%; Score 513; DB 10; Length 518;  
 Best Local Similarity 30.1%; Pred. No. 6.4e-33;  
 Matches 132; Conservative 81; Mismatches 156; Indels 70; Gaps 12;

QY 1 MGPPIGETISFFKPHRSISIGTFLQORVSRGVKFNKSGKAVYSCDDELNMFTION 60  
 DB 81 MGPIIGETCDFEPHGLYEISFVKKRMLKYGPFRINFGSNVLTLEPDIFFEVFO 140  
 QY 61 EKLFTSDPKAMHDILGKYSLLATGEIHRKLVNIIISFINLTKSKPDLHCENMISTIS 120  
 DB 141 ENKSEFVSIPFPAVKFPGKENVPLKGNHKNHKOISLQHSSEALKKMGIGIDRVTE 200  
 QY 121 ILSWKNCREVEFHKEKVFKNVSVNVLIS-IKPEDPARLYVLODFLSYMGFISLP 179  
 DB 201 HLSKKNOSGFDKAEVSEVYMAHLTPKIISNLKPEFOATL--VDNIMALSGEMFOSPK 258  
 QY 180 LRGGTNAIKVSNRNHNIATIEDM--NNAIRE--EDFLDS-----IISNED- 224  
 DB 259 L--TTLISIKVFIAR-TALQVYKIDVFTRRKASREMGDFLDTWVEBGEKEVITNEES 315  
 QY 225 -----EENHAI-RAKKGDELNWDY 245  
 DB 316 AINILFALLVAKESISSVSLAKFLAENHKAALAEKREHAAIILONRNKGAGVSEWEY 375  
 QY 246 Q-KNEFTQCVISEALRCGNIVKTVHRKATHDIKREKVFIPKGVPPFIYAVHLDLSLHE 304  
 DB 376 RHQMTFTNMYINETLRMANAPIMYRKAVNDVEKGTTPAGWIVAVIPRAVHNDIATYE 435  
 QY 305 NPEEFNMRW-----TKT-TAFGGVAVRCPGEGELGKLIQIAFFLHNLVLSYRWKIKSD 355  
 DB 436 NPLEFNMRWEGKELBSGSTFPMVFGGVCQVGAERARLQISFIHNLVTTTDFSLAQE 495  
 QY 356 EMPIAHPYVEFKRGMLEI 374  
 DB 496 SEFIRAPLPYPPKGLPIKI 514

RESULT 15  
 ID Q9FH76 PRELIMINARY; PRT; 463 AA.  
 AC Q9FH76;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE CYTOCHROME P450.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 Tabata S.,  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty pl and TAC  
 RT clones."  
 RL DNA Res. 7:31-63(2000).  
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AB020744; BAB10255.1;  
 DR InterPro; IPR001128; Cyt\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 463 AA; 52366 MW; CCD17293F53F812 CRC64;

Query Match 24.9%; Score 505.5; DB 10; Length 463;  
 Best Local Similarity 28.9%; Pred. No. 2.2e-32;  
 Matches 126; Conservative 73; Mismatches 160; Indels 77; Gaps 12;

QY 1 MGPPIGETISFFKPHRSISIGTFLQORVSRGVKFNKSGKAVYSCDDELNMFTION 60  
 DB 41 MGPPIGETISFFKPHRSISIGTFLQORVSRGVKFNKSGKAVYSCDDELNMFTION 96  
 QY 61 EKLFTSDPKAMHDILGKYSLLATGEIHRKLVNIIISFINLTKSKPDLHCENMISTIS 116  
 DB 97 KSHLKPFPKPKAKEMKQALFFHQGDYHSKRLV-----LRAFMDAIRNVPPIES 151  
 QY 117 ILSIILSKNCREVEFHKEKVFKNVSVNVLIS-IKPEDPARLYVLODFLSYMGFISLP 172  
 DB 152 IAOESLNSWDG-TQINTYOEMKTYFNVALISLGRKDEVYREDIKRCYTLDE-----KG 205  
 QY 173 FTSIPPLPGCTGYNAIKV-----RSNRNION--ATIEDMNAIREED 214  
 DB 206 YNSMPTINLPGLTFHKAARKKLAOLIANILSKRRONSSHTDLGSMED- KAGLTDEQ 264  
 QY 215 FLDSTI-----SNEDEHAI-RAKKGDELNWDY 246  
 DB 265 IADNITGVIFAARDTASVLTWILKYIADNPVLEAVTEEQMAIRKDKKEGESLTWEDTK 324  
 QY 247 KNEFTQCVISEALRCGNIVKTVHRKATHDIKREKVFIPKGVPPFIYAVHLDLSLHE 306  
 DB 325 KMLTYRIVIOETLRAATILSFTFREAVDEVEYGLIPGMYVLPDLFRINHNNADIFSDP 384  
 QY 307 FEFNPMRW-----TKT-TAFGGVAVRCPGEGELGKLIQIAFFLHNLVLSYRWKIKSD 360  
 DB 385 GKDEPFRFPAKPNFMFGSISHCPSNMLAKLEISVLHNLTKYRWSIVGPSDGIQ 444  
 QY 361 H-PYVEFKRGMLEI 375  
 DB 445 YGPFALPONGIPLALE 460

Search completed: July 29, 2002, 13:59:20  
 Job time: 107 sec



30/08:40:07 2002

us-09-995-917a-1.rspt

Page 9

1442 6005 10 0000 000000

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:26:04 ; Search time 2217.11 Seconds  
(without alignments)  
10845.015 Million cell updates/sec

Title: US-09-995-917A-3

Perfect score: 1149

Sequence: 1 atggagatgccttcattg.....caaaattcctgaagattag 1149

Scoring table: IDENTITY\_NUC

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: GenBank:\*

2: gb-ba:\*

3: gb-hg:\*

4: gb-in:\*

5: gb-om:\*

6: gb-ov:\*

7: gb-pa:\*

8: gb-ph:\*

9: gb-pl:\*

10: gb-pr:\*

11: gb-ro:\*

12: gb-sy:\*

13: gb-un:\*

14: gb-vl:\*

15: em-ba:\*

16: em-fun:\*

17: em-hum:\*

18: em-in:\*

19: em-mu:\*

20: em-om:\*

21: em-or:\*

22: em-ov:\*

23: em-pa:\*

24: em-ph:\*

25: em-pl:\*

26: em-ro:\*

27: em-sy:\*

28: em-un:\*

29: em-vl:\*

30: em-hg-hum:\*

31: em-hg-inv:\*

32: em-hg-other:\*

33: em-hgo-inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	432	37.6	100269	8	ATF18022	AL163817 Arabidops
2	174.6	15.2	1691	8	AF412114	AF412114 Arabidops
3	107.6	9.4	1398	8	AT063728	AT063728 Arabidops
4	107.6	9.4	1722	8	AT052655	AT052655 Arabidops
5	98.8	8.6	1741	8	AB035868	AB035868 Arabidops
6	97.8	8.5	1419	8	AY063722	AY063722 Arabidops
7	97.8	8.5	1608	6	AR074363	AR074363 Sequence
8	97.8	8.5	1608	8	ATCYP450R	X87367 A. thaliana
9	97.8	8.5	1610	8	AY052726	AY052726 Arabidops
10	97.8	8.5	1641	8	AT042837	AT042837 Arabidops
11	95.6	8.3	1476	8	AB066286	AB066286 Arabidops
12	90.4	7.9	2008	8	AY050980	AY050980 Arabidops
13	89.8	7.8	1681	8	AY065065	AY065065 Arabidops
14	89.2	7.8	139969	2	OSJN00012	AL066441 Oryza sat
15	89.2	7.8	169845	2	OSJN00023	AL066568 Oryza sat
16	85.6	7.4	1649	8	AE279252	AE279252 Vigna rad
17	83.8	7.3	1395	8	SL054770	U54770 Solanum lyc
18	83.6	7.3	1934	8	AB008097	AB008097 Arabidops
19	83	7.2	1461	6	AX146405	AX146405 Sequence
20	82.4	7.2	4818	8	AF044216	AF044216 Arabidops
21	82.4	7.2	84196	8	AT73A5	AL132979 Arabidops
22	80	7.0	4937	6	AR074364	AR074364 Sequence
23	80	7.0	4937	8	ATCYP450D	X87368 A. thaliana
24	80	7.0	87835	8	AB005237	AB005237 Arabidops
25	74.8	6.5	80733	8	AP001307	AP001307 Arabidops
26	67.8	5.9	1626	8	AY065073	AY065073 Arabidops
27	59.6	5.2	1740	8	AF318501	AF318501 Arabidops
28	59.2	5.2	1487	8	AF318500	AF318500 Arabidops
29	59	5.1	1503	6	AX146367	AX146367 Sequence
30	58.6	5.1	7218	8	I66494	I66494 Sequence
31	58.4	5.0	90459	8	AC012396	AC012396 Arabidops
32	58	5.0	116205	8	AC006931	AC006931 Arabidops
33	56.2	4.9	856	8	AF216313	AF216313 Helianthu
34	55.6	4.8	56287	8	AP002060	AP002060 Arabidops
35	55	4.8	160798	2	OSJN00128	AL021687 Arabidops
36	53.2	4.6	62052	8	AB020744	AB020744 Arabidops
37	51.8	4.5	1718	8	AF212991	AF212991 Cucurbita
38	51.4	4.5	90551	8	AT75K18	AL022568 Arabidops
39	51.4	4.5	96574	8	AT75K18	AL021687 Arabidops
40	51.4	4.5	198354	8	ATAP23	299708 Arabidops
41	51.4	4.5	198750	8	ATCHRIV85	AL161589 Arabidops
42	51.4	4.5	199382	8	ATCHRIV50	AL161550 Arabidops
43	51	4.4	89214	8	ATF14D17	AL135392 Arabidops
44	50.8	4.4	1494	6	AX146370	AX146370 Sequence
45	50.8	4.4	1494	8	AF318211	AF318211 Taxus cus

## ALIGNMENTS

RESULT 1

LOCUS ATF18022 100269 bp DNA linear PLN 13-APR-2000

DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone F18022 (ESSA project).

ACCESSION AL163817

VERSION AL163817

KEYWORDS

SOURCE

ORGANISM

Chale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.

1 (bases 1 to 100269)

1 Beyer, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.

2 (bases 1 to 100269)

Unpublished

EU Arabidopsis sequencing project.

Submitted (13-APR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitze 18a, D-82152 Martinsried, FRG. E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de project

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@bbsrc.ac.uk  
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.blochem.mpg.de/proj/thal/>.

## FEATURES

1. 100269

Location/Qualifiers

source

/organism="Arabidopsis thaliana"

/variety="Columbia"

/db\_xref="taxon:3702"

/chromosome="5"

gene

2255. .6118

/gene="F18022\_10"

complement(join(2255. .2376,2496. .2602,2699. .2773,

2896. .3004,3094. .3217,3419. .3493,3582. .3743,3871. .4005,

4364. .4393,4488. .4553,4723. .4759,4849. .4922,5124. .5180,

5289. .5339,5466. .5531,5640. .5775,5930. .5981,6091. .6118))

/gene="F18022\_10"

complement(2255. .2376)

/gene="F18022\_10"

/number=1

CDS

complement(join(2255. .2376,2496. .2602,2699. .2773,

2896. .3004,3094. .3217,3419. .3493,3582. .3743,3871. .4005,

4364. .4393,4488. .4553,4723. .4759,4849. .4922,5124. .5180,

5289. .5339,5466. .5531,5640. .5775,5930. .5981,6091. .6118))

/gene="F18022\_10"

/note="strong similarity to protoporphyrinogen IX oxidase, glycine max, AB025102"

/codon\_start=1

/product="protoporphyrinogen oxidase-like protein"

/protein\_id="CAB87761.1"

/db\_xref="GI:7573447"

/translation="MASCVAADHOTEAVSGKRVAVVAGVSGIAAYKLSRGILVTV

FEADSRVCGKLRSVMQNGILNDEGANTTEBAPEVGSILDDGLREKQOFRPSQKKRY

IVRNGVEVPMPLPNPLELVTSSVLSQSKFQILLPEFLMKKSSKVSADSAESVSEFP

QRHGEQEVVDYLIDPEVGTSAADPDSLMSKHSFEDLNRSFGSIIVGAIRKFAKRG

KSRPTKSSPGTKGSGSFGSGWQIILPDLCKSLSHDEINLDSKVLSTYNSGSRQ

ENMSLSGVSHNETORQNHVDAAPLCNFKEMKVMKGGQPEOLFPEINMPLSVLIT

TPTREKVRPLLEGEGVLLPSKEOKGKFKTIGLPSMMFPPRSPSDVHLITTFGIGSR

NOELAKASTDELKQVYVTSDLQRLGVECEPVSNNRYRKRAFPYLDSSDYSMEYIDK

MENDLPGFYVGNHNGSLSVGKSIASGCKADLVISTLSCSNDKKPMDSL"

complement(2377. .2495)

/gene="F18022\_10"

/number=1

Intron

complement(2496. .2602)

/gene="F18022\_10"

/number=2

exon

complement(2603. .2698)

/gene="F18022\_10"

/number=2

Intron

complement(2699. .2773)

/gene="F18022\_10"

/number=2

exon

complement(2774. .2895)

/gene="F18022\_10"

/number=3

Intron

complement(2896. .3004)

/gene="F18022\_10"

/number=4

exon

complement(3005. .3093)

/gene="F18022\_10"

/number=4

Intron

complement(3094. .3217)

/gene="F18022\_10"

/number=5

Intron

complement(3218. .3418)

/gene="F18022\_10"

/number=5

exon

complement(3419. .3493)

/gene="F18022\_10"

/number=6

Intron

complement(3494. .3581)

/gene="F18022\_10"  
/number=6  
complement(3582. .3743)  
/gene="F18022\_10"/number=7  
complement(3744. .3870)  
/gene="F18022\_10"/number=7  
complement(3871. .4005)  
/gene="F18022\_10"/number=8  
complement(4006. .4363)  
/gene="F18022\_10"/number=8  
complement(4364. .4393)  
/gene="F18022\_10"/number=9  
complement(4394. .4487)  
/gene="F18022\_10"/number=9  
complement(4488. .4553)  
/gene="F18022\_10"/number=10  
complement(4554. .4722)  
/gene="F18022\_10"/number=10  
complement(4723. .4759)  
/gene="F18022\_10"/number=11  
complement(4760. .4848)  
/gene="F18022\_10"/number=11  
complement(4849. .4922)  
/gene="F18022\_10"/number=12  
complement(4923. .5123)  
/gene="F18022\_10"/number=12  
complement(5124. .5180)  
/gene="F18022\_10"/number=13  
complement(5181. .5286)  
/gene="F18022\_10"/number=13  
complement(5289. .5339)  
/gene="F18022\_10"/number=14  
complement(5340. .5465)  
/gene="F18022\_10"/number=14  
complement(5466. .5531)  
/gene="F18022\_10"/number=15  
complement(5532. .5639)  
/gene="F18022\_10"/number=15  
complement(5640. .5775)  
/gene="F18022\_10"/number=16  
complement(5776. .5929)  
/gene="F18022\_10"/number=16  
complement(5930. .5981)  
/gene="F18022\_10"/number=17  
complement(5982. .6090)  
/gene="F18022\_10"/number=17  
complement(6091. .6118)  
/gene="F18022\_10"/number=18  
complement(6118. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"/number=18  
complement(12652. .12652)  
/gene="F18022\_20"

```
/number=1
12557. .14525
/gene="F18022_20"
Join(12557. .12652,12752. .12889,12984. .14525)
CDS
/note="Similarity to 270 kDa ankyrin G isoform, Rattus
norvegicus, EMBL:AF102552"
/codon_start=1
/product="ankyrin-like protein"
/protein_id="CAB87762.1"
/db_xref="GI:7573448"
/feature="1"
Locus="MGTDLRPOVAHVALATACRCGPDVGVTLKCGVDANSIDRL
LQSSPSLYNVDTALVAALVROVSAVRVLDQAGVYTDIMVRLGAMSDPTNGEER
RVGAGVAPPEVPTKTSRGTELRPIHTAARCSVEITIOALYFGCDLNSKNDVNTAL
SVLDCGADPEAPVKTSGTELRPIHTAARCSVEITIOALYFGCDLNSKNDVNTAL
LISTRIKHRECVKYLADGADFGVYKNGFSAVSAESNKSLEIRYLIELIRGVY
PHSSNASVSPFLYLAQAGDAEALKAUKADITLDYODEGFSAAMLAANGVEAF
RVLYAGADVAVLYLNSGDTVSLSEONGNRDIEKVMLEFALERSRMAGGAFALHC
AARGDVAVKAVLISGKGYSLDIPDGDTYPLMLAAREGHGHCYLLISCGANCAKNG
RGERLIDLATGDAEKVIRNELSRFVIEGSSVMKHTGGKCKKKGKLRMLLESSGYLS
WGSRRKRVYCKEVEIGMSORFRKNRKGGKGADEBGFVRYTTEKNEVHFVCEGELY
CAEMVKGIRLYTRETICGROTON"
12653. .12751
/gene="F18022_20"
/number=1
12752. .12889
/gene="F18022_20"
/number=2
12890. .12983
/gene="F18022_20"
/number=2
12984. .14525
/gene="F18022_20"
exon
Intron
Intron
Query Match 37.6%; Score 432; DB 8; Length 100269;
Best Local Similarity 87.1%; Pred. NO. 1.7e-91;
Matches 519; Conservative 0; Mismatches 0; Indels 77; Gaps 1;
```

```
Oy 555 tacaaagcaatgaagtgtagatccatcgataatacatcaaaagcaattatag 610
|||||
Db 63347 TCAAAACCAATTAAAGTTAGTTCATCCATCGTATATACATCAAAACGCAATATATAG 63402

RESULT 2
AF412114
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1691)
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamuya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis CDNA clones
Unpublished
2 (bases 1 to 1691)
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamuya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (21-AUG-2001) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.
Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
source
1..1691
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAF106-76-F19(R13551)"
/note="ecotype: Columbia"
1..28
29..1570
/note="steroid 22-alpha-hydroxylase (DWF4)"
/codon_start=1
/product="AT3g50660/73A5_40"
```

```

/Protein_id="AAL06567.1"
/db_xref="GI:15724348"
/translation="MFEFHHLLPLLLPSLLPLLLKRRNRKRNLPFGKS
CMPYIGERTIGYLTATATTLGDPFMOOHYSKTKYTRSNLGEPTTVSADAGNRLILO
NGRLFECSYPSRISIGILGKWSMLVVGDMHDMQSIISNFIISHARLITLLKDYERH
TLEVLDSWQNSIFSAODEAKKEFTFNLAKHIMSDPGEETEOJKEVITPMKGVSA
APLNIPGTAYHKAOLSRATILKEIERKMERLIDJKEEOEEVYKDEEAMSKSDH
VKORITDLDLGMVLSKNSLSTEOILDLISLTFAGHETSVAVALIPILOACPKAV
BELREHEIARAKKEGSELMNDPYKKRMOCVINTLILGVVPRILKRAIKDV
RYKGYDISGKWKVLPVISAHVHLDNSRYQPNFNNRMOQNNNGSSSGSFTWGN
NMPFGGGRILCAGSELAEKAVFIHLVLKFNWELAEKDPFAFPEVDFPGLPIR
VSRIL"
3'UTR 1571..1691
BASE COUNT 524 a 286 c 363 g 518 t
ORIGIN

Query Match 15.2%; Score 174.6; DB 8; Length 1691;
Best Local Similarity 53.6%; Pred. No. 8.1e-31;
Matches 363; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

QY 4 gtagtgccttcatttggaagaactatttcttctcaaacctatagatcagactccac 63
DB 161 GGTGGGCATTTCTTGTAAGAACCATCGGTTATCTTAAACCGTACACGCCACACACATC 220
QY 64 gttacattcttgcaacaacgtgttcacgtatggaagaagtgtcagaatcaatata 123
DB 221 GGTGATTCATGCACACACATGCTCTCCCAAGTATGATATATATGATCGAACTTGTT 280
QY 124 ggtggaagaagcagtagtctcatgtgaccaagaactcaacatgtctactctcaaa 183
DB 281 GGAGAACCAACGATCGTACAGTGTGCTGACGCTTAATAGATTCATATTAACAGAA 340
QY 184 gggagagtgttactcagattataccaagaagatcagactctcgcgaatattcc 243
DB 341 GGAAGGCTCTTGAATGATTAATCTTAAGAATAGTGGTGGGTTTGGGAATGGTGG 400
QY 244 ctctctatgacacccgagaagaattcacagaagaactaaaaatgtatattagctcac 303
DB 401 ATGCTTGTCTTGTTGGTACATGATAGATAGATAGAGAAATATCTCGCTTAACCTCTTA 460
QY 304 aatctacaagaatcgaaccttcttcttctcactgcgcagaagaacctctctatcga 363
DB 461 AGCTACAGCACGCTTAGAATCTTCTACTTTAAAGATGTTGAGAGACATCTTGTGTT 520
QY 364 ctaaaatcagaaaatattgcgagaagtcgaattccataagaagttaaaagtattac 423
DB 521 CTTGATTTCTTGGCACACAACTATTTCTCTGCTCAAGACGAGGCCAANAAGTTTACG 580
QY 424 ctcaagtgtatgttaaaccaactcttgagcatcaagccagaagaaccagaagacttat 483
DB 581 TTTAATCTAATGGCGAAGCATATTAATAGATAGATCTCGAGAAAGAAACAGAGCAA 640
QY 484 gtaattgcaagaattttatctatatgaagggttattctccttaccataccgctcca 543
DB 641 TTAAGAAAGAAAGATAGTAACTTTCATGAAGAGAGTTGCTCTGCTCTTAATCTACCA 700
QY 544 ggaacgggtatatacaaacgcaatlaaggttagatccaatcgaatatatatacaacga 603
DB 701 GGAACGCTTATCATTAAGCTCTTCAAGTACAGGCAACGATATTGAGTTCAATGAGAG 760
QY 604 attatagaagacatgataataagaagaagaagatttctcgatctcgataattcgc 663
DB 761 AAAATGGAAGAGAGAAATTTGGATATCAAGAGAAAGATCAAGAGACAGAAAGAGTAAA 820
QY 664 aatgaagatgaagaaca 680
DB 821 ACAGAGAGATGAAGACGA 837
LOCUS 3 AY063728 1398 bp mRNA linear PLN 05-DEC-2001

```

```

DEFINITION
Arabidopsis thaliana AT3g30180/T20F20_6 mRNA, complete cds.
ACCESSION
AY063728
VERSION
AY063728.1 GI:17380629
KEYWORDS
FTL_CDNA.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
Rosidae; eurosid II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1398)
REFERENCE
Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1398)
Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
Direct Submission
Submitted (15-NOV-2001) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN
Arabidopsis Full-Length cDNA") : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamliya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs : Kim,C.J., Chen,H.,
Cheuk,R., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Chang,E.,
Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A.,
and Ecker,J.R.
Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
source
location/Qualifiers
1..1398
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U13734"
/note="this clone is in pUNT 51
ecotype: Columbia"
1..1398
/note="putative cytochrome P450 homolog"
/codon_start=1
/product="AT3g30180/T20F20_6"
/protein_id="AAL36078.1"
/db_xref="GI:17380630"
/translation="MGIMMIMILGLIIVICLTALLRNNOHRYSKGIPPTMGPIF
GETTEFGDPMKORLRYGSEFKSHIGOPTIVSDAENRILNNEKRGVAGY
POMSDIITGONIAAIVHGPSRIMAGSILSLISPMMDHLLPRTIDPMRYLIGAMD
LEIVDIOETKTHMAFLSSLDQIAETLKPBEVETFEFKLVGTLSPIDPIGNIR
SGVQARNIIDRLTELMOERKESGETFDMLGMLKEDNNRLILDTKISPIROOVVILY
SGYEVSTSMALKYLDHDKALEELREHAIERKRDEPLLDIDIKSMKFTFRAV
IETSRLATIVGVYRKTTHLELNGYLIPGWRIVYVYTRINVDTSLEYEDPMFNPV

```

BASE COUNT	448 a	258 c	304 g	388 t
ORIGIN	RMHMKSLSEKSYFLIFGGGVRNLCKGKRELGISVSSFLHYFTYKTRMEENEDKLMTFF RVSAFKGTHLKCSPY"			
Query Match	9.4%	Score 107.6:	DB 8;	Length 1398;
Best Local Similarity	56.1%;	Pred. No. 4.8e-15;		
Matches 240;	Conservative 0;	Mismatches 164;	Indels 24;	Gaps 1;
Oy	675	agaagaatgcacccatctagaagcccaagaagaaggagatggggaactcttgaaatctggagaatata	734	
Db	915	AGAAATATTGGCTTAATAGAGGAGAGAAACGACCTACGACCGCTTCTCGACGATAT	974	
Oy	735	tcaagaagaatggaattcactcaaatgctgtaattctctaggcaactacgattgylaataatcgt	794	
Db	975	TAAATCGATGAAATTTCTACTCCGACGCTGATATCTTTGAGACATCAAGATTGGCAAGCATTTGT	1034	
Oy	795	caagaatgacatagaaaaagctactcaatgatataataatcaagaataatgtaattccaaa	854	
Db	1035	TAAATGCTCTCCCTTAGGAAAAACACTACACGACCTAAGAACTCAACGGTTTATTCCTCCAAA	1094	
Oy	855	gggggggaaggagtgcttccaatctctcaagcagatataatcttgatccctctcttcacgaaa	914	
Db	1095	AGGTTGGGAATTTTCGTATACACAAGAGAGATTACTATGATACATCTCTTATGAGAAG	1154	
Oy	915	tcccttgaatttaattcccatgagatggacaacaaacga-----	952	
Db	1155	TCCATATGATCTTTAAACCCATGGAGATGCGATGGAAGAAAGCCTTAGAATCAAGAGCTATT	1214	
Oy	953	--cgagcgtttggaggaggtaaggatgctgctctggtggtgaacttggcaagctccaaat	1010	
Db	1215	CTTACTCTTTGGAGGTGAGGAGTTAGAGCTTTGCCCTGAAAGAACTAGGAATCTCGGAAGT	1274	
Oy	1011	tgtctctctctctcaatctatctgtctctctccataggtgggaaaaaaagtcagatggaat	1070	
Db	1275	CTCAAGCTCTCTTCACTACTTGTTCACAAAATATGATGGGAAGAAATGGAGAAACAA	1334	
Oy	1071	gccaatcg 1078		
Db	1335	ATTAAATGG 1342		
RESULT 4				
AY052655	1722 bp	mrna	linear	PLN 05-SEP-2001
LOCUS				
DEFINITION	Arabidopsis thaliana At3g30180/T720F20_6 mrna, complete cds.			
ACCESSION	AY052655			
VERSION	AY052655.1			
KEYWORDS	GI:15450574			
SOURCE	FLI CDNA.			
ORGANISM	thale cress.			
REFERENCE	Arabidopsis thaliana			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
1 (bases 1 to 1722)				
Chauk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamlya,A., Karlin-Neumann,G., Kawasaki,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Tsurumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
Arabidopsis CDNA clones				
Unpublished				
2 (bases 1 to 1722)				
Chauk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamlya,A., Karlin-Neumann,G., Kawasaki,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Tsurumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
Journal				
REFERENCE				
AUTHORS				

Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

Direct Submission  
Submitted (17-AUG-2001) Salk Institute Genomic Analysis Laboratory (Simgal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA.') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koeseama,E., Meyers,M.C., Shinn,P., Banji,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1. .1722

/organism="Arabidopsis thaliana"

/db.xref="taxon:3702"

/chromosome="3"

/clone="RAFL06-74-K19(R1374)"

/note="ecotype: Columbia"

1. .91

92. .1489

/note="putative cytochrome P450 homolog"

/codon\_start=1

/product="AT3g30180/T20F20.6"

/protein\_id="AAK96559.1"

/db\_xref="GI:15450575"

/translation="MGIMMMILGLVIVICLTALRMNOMRKKKLPGTMGPPIFGETTEFLKQGPDPFKNRRLRGSEFFKHSILISPTFKMDHLPRIDPMKRYLCGWD POSMDLFLGTCNTIAAVHGPSRHLRGLSLISPTFKMDHLPRIDPMKRYLCGWD LEYDIDIEKTRHMAFLSLSLQIAETLAKPEVEYRTFFPLVGTSLSPIDIPGNR SGVQARNNIDRLTELMOERKESGETFDIMGYMKKEDNRYLTDKEIRDOVVTILY SGVETSTTSMALAKYLHDHPRLAEELIRREHLIRERKRDPEPLTLDIKSMKPTRAV IFEFSRLATVNGVLRKRTYDLDLELNGVLRGWRIVYTRREINDGSLYDMPENFW RMKRSLSKRSYFLLEGGGVRALCPGKELGISVSSFLHYTYTKRMEDNEDKLWPP RVSAPKGYHLKSPY"

RVSAPKGYHLKSPY"

1490. .1722

3'UTR 585 a 301 c 351 g 485 t

BASE COUNT

ORIGIN

Query Match 9.4%, Score 107.6, DB 8, Length 1722;  
Best Local Similarity 56.1%; Pred. No. 4.7e-15;  
Matches 240; Conservative 0; Mismatches 164; Indels 24; Gaps 1;

675 agaagatcagccattagaagccaagaagaaggatgggaaccttggaaattgggaagelta 734  
|||||

Db 1006 AGAACAATTTGGTAAAGGAGAGAAAGCAACGACTGACGACCGCTCTCGACGATAT 1065  
|||||

QY 735 lcaagaagatggaactcacatcaaaatgtgtatcttctgaaggcaactacgaatggtlaataatcgt 794  
|||||

Db 1066 TAAATCGATGGAATTTACTCTCGAGCGTGATATCTTTGACGACATCAACGATTGGCAGCATTTGCT 1125  
|||||

QY 795 caaagctctacatagaaaagctacatcatgatataatcaatcaagaataatgtgatccaa 854  
|||||

Db 1126 TAATGCTGCTTTAGGAAACACTACACGACTTACAAACGGTTATTTAATCCCAA 1185  
|||||

QY 855 ggggtggaagtggttcccaactctcaagcgagtaactcttgatccctctctcaagaaga 914  
|||||

Db 1186 AGCTTGGAATTTACGTTACACAAAGAGATTAACTATGATCATATCTTTATGAAGA 1245  
|||||

QY	915	ttcttctgattatcccatcagataggcccaaaaga-----	952
Db	1246	TCCATGATCTTAAACCCATGGAGATGAGTGAAGAAAGACCTTAGAATCAAGAAGCTATT	1305
QY	953	--cgagctttgagagaggaaggaaggaagcagtcctgctgctgagactgtgcaagctccaat	1010
Db	1306	CTTACTCTTTGGAGGTGGAGTTAGCGTTTCCCTGGAAAGCAACTGAGATCTCGGAAT	1365
QY	1011	tgcttcttctcattcattcctgtgctctcattataggctggaataaagtcagatgaaat	1070
Db	1366	CTCAACCTTCCTCTCACTACTTGTTCACAAAATATAGATGGAGAAAGAAATGAGAAACAA	1425
QY	1071	gccaatcg 1078	
Db	1426	ATTAAATGG 1433	
RESULT	5		
LOCUS	AB035868		
DEFINITION	AB035868	1741 bp mRNA linear	PLN 28-JUL-2001
ACCESSION	AB035868		
VERSION	AB035868.1	GI:14475585	
KEYWORDS	brassinosteroid-6-oxidase; Brassinosteroid-6-Oxidase (P450).		
SOURCE	Arabidopsis thaliana (strain:columbia) cDNA to mRNA.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 1741)		
TITLE	Shimada, Y., Fujikawa, S., Miyachi, N., Kuehito, M., Takatsuto, S., Nomura, T., Yokota, T., Kamiya, Y., Bishop, G. U. and Yoshida, S.		
JOURNAL	Brassinosteroid-6-oxidases from arabidopsis and tomato catalyze multiple c-6 oxidations in brassinosteroid biosynthesis		
MEDLINE	Plant Physiology. 126 (2), 770-779 (2001)		
REFERENCE	21295570		
AUTHORS	2 (bases 1 to 1741)		
TITLE	Shimada, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (15-DEC-1999) Yukihisa Shimada, RIKEN, Plant Functions;		
FEATURES	2-1 Hiroseawa, Wako, Saitama 351-0198, Japan		
source	(E-mail:shimada@postman.riken.go.jp, Tel:048-467-9525)		
gene	Location/Qualifiers		
CDS	1..1741		
	/organism="Arabidopsis thaliana"		
	/strain="columbia"		
	/db_xref="taxon:3702"		
	112..1509		
	/gene="BR6ox1"		
	112..1509		
	/gene="BR6ox1"		
	/note="P450"		
	/codon_start=1		
	/product="brassinosteroid-6-oxidase"		
	/protein_id="BAB60858.1"		
	/db_xref="GI:14475586"		
	/translation="MGAMVMMGLLILIVSLCALLRNOMRYTKNGLPPTGMPLEIFR		
	GEMPEFLKQGNPMRNOILRYGSEFKSLILCPTLISMDSVNRYTILKNESKGLVPG		
	POSMIDLITGCTNMAAYHSSSRILKRGSLISLTSMADHILPYVDHPRSYLDOONN		
	LEVIDIDQTKHMAFLSLVITIGALNRPFEVEERTKAEFLVSTLSPIDLPTNFR		
	CGIARNNIDILRLKLMQERDSDGTFEDMILGYLMKKGNGYPLTDEIRDOVYTIY		
	SGEYVSTTSMALAKYLHDHPKALQELDMELADPERRKROEPLGLEBVKSMKPTRAV		
	IYERSLATVINGVLRKTRDLEINGYILPGWRIYVYTRINDANIELDPLFENP		
	RWKKSLSEONSCVFGGGRCLCPKKEGLIVEISSFLHYTRIRWELIGDELAVTFEP		
	RVPARKGRHLISPY"		
BASE COUNT	544 a	317 c	371 g 509 t
ORIGIN			
Query Match	8.6%	Score 98.8	DB 8; Length 1741;
Best Local Similarity	52.7%	Pred. No. 5.5e-13;	
Matches 257; Conservative	0;	Miss. 5.5e-13;	Indels 24; Gaps 1;

QY	676	gaacatgcgcgaccattagagccaagaaggsggttgsgaacctttagaatgggaagtatt	735
Db	1027	GAGCATTTGGCATTACAGGGAAGAAGAACACAGAGCAACAACCTCGTCTTAGGACGTG	1086
QY	736	cagaagatgaaatcaactcaatcgaatgltgatcttcctcgagccactcagatgltgtaatacgtc	795
Db	1087	AAGTCATTAAGATTCACCTCGAGCTGTGATTTATGAGACATAAAGTATGGCAACGATCGTT	1146
QY	796	aagaactgtcatagaaaagacctactcatgatataatcaaaagaatatgatlgtccaaag	855
Db	1147	AATGGGGTCCCTAAGGAAGAAACTCTGTCGCTTGGAATCAACAGGTTATTAATCCAAAA	1206
QY	856	ggatgaaagtgltttcccaatcttccacaagagfacactctgtacccctccctatgaatat	915
Db	1207	GGATGAGAAATTTANSTATACACGAGGGGAATTAATTACGATGCCAAATCTTATATACAGNC	1266
QY	916	cccttgaattaatccatccatgagatgagaccaaagca-----	952
Db	1267	CCATTGATCTTTAATCCATGAGATGAGATGAGAAGAGCGCTTGAGTCCAAAATCATGC	1326
QY	953	-cgcggtttggagggaggaataaggatgctgtgtgtgaactctggcaagctccaatt	1011
Db	1327	TTTGTGTTTGGAGGGGAGGAGCGCTTGTCCGTGTAAAGAACTGAGGCTTTCAGATC	1386
QY	1012	gcttcttcctcatcatatgcttctctcctaatagttgtaaataaagtcagatgaatg	1071
Db	1387	TGGAGCTTCTCCATTACTTTGTATGAGAGATACAGATGGAGAGAAATAGAGGGGATGAA	1446
QY	1072	ccaatgcgcacccttaacgtlgaagttaagaagagcatgcttltyagatagaagccaaca	1131
Db	1447	TTAATGGTWTTCACAGAGTTTTTTCACCACCAAAGCCTTCATCTAGGATTTTCACCCTTAC	1506
QY	1132	aaattcct	1139
Db	1507	TTTTT	1514
RESULT	6		
LOCUS	AY063722	1419 bp	mRNA linear PLN 05-DEC-2001
DEFINITION	Arabidopsis thaliana AF590569/MJ03_9 mRNA, complete cds.		
ACCESSION	AY063722		
VERSION	AY063722.1	GI:17380617	
KEYWORDS	FLI, CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banb,J., Bowers,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Oondera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1419)		
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banb,J., Bowers,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Oondera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-Nov-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for		





```

/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/map="14.3 cM"
/clone="C204"
/clone.lib="lambda gt10"
48..1466
/gene="CPD"
/gene="CPD"
48..1466
/gene="CPD"
/function="loss of gene function results in dwarf phenotype"
/note="cytochrome P450"
/codon_start=1
/product="CYP90 protein"
/protein_id="CA60793.1"
/db_xref="GI:853719"
/translation="MAFTAFLLLSIAAGFLLLRTRYRMGLPPGSLGPIIGET
FOLIGAYKTENPEPFIDERVARYGVFMTHLREPTIFSNADPTNRFVLQNEKLEEC
STPASTCNLIGHSLILMKSLKRMHSLTMSFANSIIKDHMLDIDRLVRNLDWS
SSRVLMEBAKKITFELTYKOLMSPDPMSESLREKVELIVIEGFSLEPLFSTYR
KAIOARKYAEALTYVMKRREBEERGAERKKMDLAAILAADGFDEDELVDLVAL
VAGETSTIMTLAVKFLIETPLALQLEBEHKIRAKMSDSYLSMSYKSPFTQC
VNETLRVANIIGGVFRAMTVEIKGYKIPKGMKFSSFRAYHLDPNHFKDARTNP
WRQNSVTTGSSNFTPPGGGRLCPGIELARVALSVFLHRLVTGFSWVPAEDKLY
FPFTTRKRYPIFYKRDFAF"

BASE COUNT      382 a      374 c      374 g      478 t
ORIGIN

Query Match      8.5%; Score 97.8; DB 8; Length 1608;
Best Local Similarity 58.9%; Pred. No. 9.5e-13;
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 671 atgaagaacatgcagccattagagccaagaaggatggygaactttgaattggaag 730
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 AGGAAGAGCATGAAAGATTAGGCAATGAAGACTGATTCTATAGCTTGAAGAGAGTG 1020
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 731 attacagaagaatggaactcaatcattgattcttgagggacatacagatggtgataa 790
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 ATTACAGATCAATGCCATTGACCAATGTGTGTTATGAGACGCTACGAGTGGCTAACA 1080
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 791 tcgtcagaactgtcacataagaagctctcatatataatcaagaataatgtgttc 850
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 TCATCGGCGGTGTTTCAGACGTCAATGACGATGTTGAGATCAAGTTATTAATAATTC 1140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 851 caaagggtggaaggtgttcacaaattccacagcagatcacattgacccctctcatg 910
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 CAAAAGGTGGAAGATTCTTCATCGTTTAGAGGGTTTCATTAGACCAACCAACTTCA 1200
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 911 aaatcctttgaattaatccatgagatgagaccaaaacgacgg 955
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 AAGATGCTCGCACCATTCAACCTTGAGATGCGACAGCAACTCGG 1245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AY052726 Arabidopsis thaliana 1610 bp mRNA linear PLN 05-SEP-2001
DEFINITION Arabidopsis thaliana AT5g05690/MJ3_9 mRNA, complete cds.
ACCESSION AY052726
VERSION AY052726.1 GI:15450716
KEYWORDS FLI_CDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
            1 (bases 1 to 1610)
REFERENCE Shin,P., Chen,H., Cheuk,R., Kim,C.-J., Koesema,E., Meyers,M.C.,
AUTHORS Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,

```

```

Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1610)
Shin,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C.,
Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (20-AUG-2001) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFL cDNAs (RFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PECC (SSP) Consortium members carried out the
sequencing and annotation of the RFL cDNAs: Shin,P., Chen,H.,
Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Shin,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
    source
        /organism="Arabidopsis thaliana"
        /db_xref="taxon:3702"
        /chromosome="5"
        /clone="RAF106-74-M20(R13805)"
        /note="ecotype: Columbia"
        1..119
        20..1438
        /note="cytochrome P450 90A1 (sp1042569)"
        /codon_start=1
        /product="AT5g05690/MJ3_9"
        /protein_id="AK96630.1"
        /db_xref="GI:15450717"
        /translation="MAFTAFLLLSIAAGFLLLRTRYRMGLPPGSLGPIIGET
        FOLIGAYKTENPEPFIDERVARYGVFMTHLREPTIFSNADPTNRFVLQNEKLEEC
        STPASTCNLIGHSLILMKSLKRMHSLTMSFANSIIKDHMLDIDRLVRNLDWS
        SSRVLMEBAKKITFELTYKOLMSPDPMSESLREKVELIVIEGFSLEPLFSTYR
        KAIOARKYAEALTYVMKRREBEERGAERKKMDLAAILAADGFDEDELVDLVAL
        VAGETSTIMTLAVKFLIETPLALQLEBEHKIRAKMSDSYLSMSYKSPFTQC
        VNETLRVANIIGGVFRAMTVEIKGYKIPKGMKFSSFRAYHLDPNHFKDARTNP
        WRQNSVTTGSSNFTPPGGGRLCPGIELARVALSVFLHRLVTGFSWVPAEDKLY
        FPFTTRKRYPIFYKRDFAF"

BASE COUNT      399 a      362 c      376 g      473 t
ORIGIN

Query Match      8.5%; Score 97.8; DB 8; Length 1610;
Best Local Similarity 58.9%; Pred. No. 9.5e-13;
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 671 atgaagaacatgcagccattagagccaagaaggatggygaactttgaattggaag 730
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 933 AGGAAGACATGAAGAAGATTAGGCAATGAGATGATTGCTATAGTCTTAAATGAGAGTG 992  
 Qy 731 attatcagaagatcgaatcactcaatgctgattctcgaagcactgaatgtaata 790  
 Db 993 ATTACAAAGTCATGCCATTCACACAAATGCTGTGTTAATGAGACGCTACAGTGGCTAAC 1052  
 Qy 791 tgcgcagactgacatagaagaagctaccatcagatataaataaagaatgtaac 850  
 Db 1053 TCATCGCGCGGTGTTTTCACACGCTGCAATGACGATTTAGATCAAAAGCTTAAATTC 1112  
 Qy 851 caaagggtggaaggtgttccaatctcagacagcagatcattgacctctcctcag 910  
 Db 1113 CAAAGGCGGGAAGATTCCTGATCTGTAGAGCGGCTCATATTAGACCAACCACTTCA 1172  
 Qy 911 aaatcccttgaattaatcccatgagatgagacaaagacagcg 955  
 Db 1173 AGATGCTCGCACTTCAACCCCTTGAGATGGCAGCAACTCGG 1217  
 RESULT 10  
 AY042837 1641 bp mRNA linear PLN 04-JUL-2001  
 LOCUS Arabidopsis thaliana cytochrome P450 90A1 (MJJ3.9) mRNA, complete  
 DEFINITION cds.  
 ACCESSION AY042837  
 VERSION AY042837.1 GI:14596098  
 KEYWORDS P450,CDNA.  
 SOURCE FLI,CDNA.  
 ORGANISM thale cress.  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1 (bases 1 to 1641)  
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,  
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,  
 Chauk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,  
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,  
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,  
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submision  
 Submitted (25-JUN-2001) DNA Sequencing and Technology Center  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
 Arabidopsis Full-Length cDNA'); Seki,M., Narusaka,M., Ishida,J.,  
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.  
 The Sait, Stanford, PGEC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Nguyen,M.,  
 Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,  
 Bowser,L., Jones,T., Banh,J., Chen,H., Chauk,R., Chung,M.K.,  
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,  
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
 (SSP/Stanford) contributed equally to this work as PIs.  
 FEATURES  
 source location/Qualifiers  
 1..1641  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /chromosome="V"  
 /clone="RAFL04-19-J21"  
 /note="This clone is in pluscript  
 ecotype: Columbia"  
 1..1641  
 /gene="MJJ3.9"  
 20..1438  
 /gene="MJJ3.9"  
 CDS

/codon\_start=1  
 /product="cytochrome P450 90A1"  
 /protein\_id="AAK68777.1"  
 /db\_xref="GI:14596098"  
 /translation="MAFTAFLLLSSTNAGFLLLRRTRRRMGLPGSGCLPIGET  
 /FOLIGAVRTENPEPFIIDERVARYGSVMTFLIEGPTIFSADEPTNRVLONEGFLPIC  
 SYPASINELKSHLSLLMKSGSLKRMHSLTMSFANSSIIDHMLDIDRLVRFMLDSM  
 SSRVLMNEAKKITFELVLKQLMSFDGEMSESLRKEVLIVIEGFSLPIPLFSTTVR  
 KAIGARRVVAALVVMKKRREEDEKMDLVAALADDFSEEDIVDFVALL  
 VAGETSTTMTLVKRTLETPPLALQIKERHEKIRAMKSDSYSLMSDYKSMPEFCO  
 VNNETLRVANTIGVFRPRAAMDVEIKKIKRKGVKSSPFAVHLDPNHFDARTFP  
 WRWOSNVSTGPSNVFPFGGPRICPEYELARVALSVFLHRLVTGFSWPAEDDKLV  
 FFPTRQKRPPIEVKRRDFAT"  
 BASE COUNT 402 a 371 c 380 g 488 t  
 ORIGIN  
 Query Match 8.5%; Score 97.8; DB 8; Length 1641;  
 Best Local Similarity 58.9%; Pred. No. 9,5e-13;  
 Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
 Qy 671 atgaagaacatgcagccattagaccagaaggagatgggaacttgaattgggaag 730  
 Db 933 AGGAAGACATGAAGAAGATTAGGCAATGAGATGATTGCTATAGTCTTGAATGAGAGTG 992  
 Qy 731 attatcagaagatcgaatcactcaatgctgattctcgaagcactgaatgtaata 790  
 Db 993 ATTACAAAGTCATGCCATTCACACAAATGCTGTGTTAATGAGACGCTACAGTGGCTAAC 1052  
 Qy 791 tgcgcagactgacatagaagaagctaccatcagatataaataaagaatgtaac 850  
 Db 1053 TCATCGCGCGGTGTTTTCACACGCTGCAATGACGATTTAGATCAAAAGCTTAAATTC 1112  
 Qy 851 caaagggtggaaggtgttccaatctcagacagcagatcattgacctctcctcag 910  
 Db 1113 CAAAGGCGGGAAGATTCCTGATCTGTAGAGCGGCTCATATTAGACCAACCACTTCA 1172  
 Qy 911 aaatcccttgaattaatcccatgagatgagacaaagacagcg 955  
 Db 1173 AGATGCTCGCACTTCAACCCCTTGAGATGGCAGCAACTCGG 1217  
 RESULT 11  
 AB066286 1476 bp mRNA linear PLN 20-JUL-2001  
 LOCUS Arabidopsis thaliana mRNA for CYP90D, complete cds.  
 DEFINITION AB066286  
 ACCESSION AB066286.1 GI:14971016  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (cultivar:Columbia) cDNA to mRNA,  
 clone:PGMD26.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1 (bases 1 to 1476)  
 Shimada,Y.  
 P450 gene repressed by brassinosteroid  
 unpublished  
 2 (bases 1 to 1476)  
 Shimada,Y.  
 Direct Submision  
 Submitted (18-JUL-2001) Yukihisa Shimada, RIKEN, Plant Science  
 Center, 2-1, Hirosewa, Wako, Saitama 351-0198, Japan  
 (E-mail:shimada@postman.riken.go.jp, Tel:81-48-467-9529,  
 Fax:81-48-462-4959)  
 FEATURES  
 source location/Qualifiers  
 1..1476  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="PGMD26"  
 /note="P450"

Query Match	8.3%	Score 95.6	DB 8	Length 1476
Best Local Similarity	55.2%	Pred. NO. 3.1e-12		
Matches 222	Conservative 0	Mismatches 159	Indels 21	Gaps 1
QY 697	aagaaagggaatggygaaccttttgtaattggaagaattacagaagatgaaatcactcaa	756		
Db 1021	AAGGAATTGACAGGAGACCACCTATATTGGAATGACTACTGTGCTTACCTTTACACAA	1080		
QY 757	tgtygatcttcttgaggcaactcgcgatgfygtaatacgcgaagctgtacatagaaga	816		
Db 1081	AAGGATTTTACAGAGACACTGAGATGGGAAATGTTAATTGAGATGATGAGAAAGCGC	1140		
QY 817	acctgatctaatcaatcaagaataatgtaattccaagaagggtggaagtggtttccaatc	876		
Db 1141	ATGAAAGATGTTGAAATTAAGATATGTGCTATACCAAAAGATGCTGTTCTTGCCCTAT	1200		
QY 877	ttcacgaagatcaactctgataccctcttcataagaataaccttttgatattaatccatg	936		
Db 1201	CTCAGATCACTGTTTCATCTTGATTAACCTTATATGATGCTCCCTCAAAATTAATCCCTGG	1260		
QY 937	agatggagccaaaag-----acggcglttggaaggaaggaag 975			
Db 1261	AGATGGCAGAAAGGACATGACACAGATGATGTTTCAGTCCCTTTGGAGGTGGCAGAGA	1320		
QY 976	gtatgtcctgtgtgtaactgtgcaagctccaatgtcttcctcactcatcattgctc	1035		
Db 1321	TTTGTCCTGGTCTCGATTTGGTGCCTGCTTTGAAACTTCAGTTTCTCTCACCATCTTGTC	1380		
QY 1036	ctctccataggtggaataaagtcagatgaaatgcgaatc 1077			
Db 1381	ACTGCTTTCAGATGAGATGACAGAAAGACACATCATTAAC 1422			
RESULT 12				
LOCUS	AY050980	2008 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana putative cytochrome P450 protein (At4g19230)			
ACCESSION	AY050980			
VERSION	AY050980.1			
KEYWORDS	FLI-CDNA.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
AUTHORS	Yamada, K., Liu, S. X., Pham, P. K., Banh, J., Banno, F., Dale, J. M., Goldsmith, A. D., Jiang, P. X., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,			

TITLE	JOURNAL	REFERENCE	AUTHORS
Full length cDNA of gene AT4g19230 (GI:7268718)	Unpublished	2 (bases 1 to 2008)	Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission	Submitted (07-AUG-2001)	Plant Gene Expression Center, 800 Buchanan Street, Albany CA 94710, USA	
COMMENT	JOURNAL		
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Ngwen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.			
FEATURES	source		
gene	5' UTR		
CDS			
misc_difference	3' UTR		
gene="AT4g19230"	misc_difference	1772..1773	gene="AT4g19230"

		/note="compared to genomic sequence"	
		/replace="ttg"	
BASE COUNT	627 a 398 c 387 g 596 t		
ORIGIN			
Query Match	7.9%; Score 90.4; DB 8; Length 2008;		
Best Local Similarity	54.2%; Pred. No. 5.1e-11;		
Matches 218; Conservative	0; Mismatches 166; Indels 18; Gaps 1;		
Qy	672 tgaagaacatcagccatagagccaagaaggagatgggaacttctgaattgggaaga 731		
Db	1051 TGAAGAAATATGCGAATAGGAAGACAAAGAGAGTCTTAACCTTGGGAGA 1110		
Qy	732 ttatcagaagatggaattcactcaatgctgcttctcagagcactgagtggtcaat 791		
Db	1111 TACAAAGAAAGATGCCATTACTTCAAGAGTTTATCAAGAACATTAAAGAGTGCCTCAAT 1170		
Qy	792 cgtcaagactgtacatagaaaagctactcatatataataatcaagaatatgtattcc 851		
Db	1171 CTTATCTTTCACATTCACAGAGAGCTGCGAAGATGTGCAATACGAAGATATTGTATACC 1230		
Qy	852 aaaggaggaggaagtggttctcaatcttccagcagatcatctgttccctctcatga 911		
Db	1231 TAAAGGATGGAAGGTGTACCCCTATTCAGAAACATCATCTGATGATATTTTTC 1290		
Qy	912 aaatcccttgaatttaattcccatgagat-----gagccaagaagac 953		
Db	1291 TAATCGGGGAAATTTGATTCATCAAGATTCGAGTGCCTCCAAACCAATACGTTCAT 1350		
Qy	954 ggcgcttggagaggagtaaggatgtctctggttggtgaacttggcaagctccaaattgc 1013		
Db	1351 GCCATTGGCAATGGAACCCACTCGTCTGCTGGAATGACCAAGCATTGAGATGTGC 1410		
Qy	1014 ttctctctcaatctctgtctctctctcatatagtggtgaaaat 1055		
Db	1411 TATTATGATTCATCATCTACCCACCAAGTACAGTTGTCAT 1452		
RESULT 13			
AY065065	1681 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana AT5G45340/K9E15_12 mRNA, complete cds.		
ACCESSION	AY065065		
VERSION	AY065065.1 GI:18086489		
KEYWORDS	FLI cDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1 (bases 1 to 1681)		
AUTHORS	Chouk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamliya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1681)		
AUTHORS	Chouk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamliya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-NOV-2001) Salk Institute Genomic Analysis Laboratory		

COMMENT	(SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN Arabidopsis Full-length cDNA": Seki, M., Narusaka, M., Ishida, J., Saitou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.			
	The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.			
	Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.			
FEATURES	Location/Qualifiers			
source	1..1681			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
	/chromosome="5"			
	/clone="RAFL09-13-J06(R21133)"			
	/note="ecotype: Columbia"			
	1..82			
5'UTR	83..1474			
CDS	/note="cytochrome P450"			
	/codon_start=1			
	/product="AT5g45340/K9E15_12"			
	/protein_id="AAL57698.1"			
	/db_xref="gi:18086490"			
	/translation="MDESGFLTHTLSAALFLCLRLFRNGVRSSTKLPDPGTMGYP YVERFDLISODPNVFFAKORRTGVSFTKTVLQPCVMISPEAKRVLTKSLFK PTFPAKSRMLDKQAIFFHQGDYHSKRLKLVLRAMPDAIRNMVPHISIAQESLSNW DGTOLNTQOEKMTYTFYFNVALISILEKDEVYREDLRQCYLLEKGVNSMPNLPGTLF HKAMKARKELOILANILSKRRONSSHTDLGSPMEDKAGLTDOINDNLTIGVFVA RDTTASVLTWTLKADNPVLEAVTEEMOILIRDKKESGLTMDPTKMPDLTYRVIO ETLRAATILSTPREAVDEVYEGVLLIPKQKVLPLPFRNHNNADIFSDPKRFPDSRP EVARPNTEMPFGSGTSHSCPGNEKLKELISVLIIHLITTKYRMSTVPSDGIQYGFPLD PONGPLALERKP"			
	1475..1681			
BASE COUNT	556 a	353 c	323 g	449 t
ORIGIN				
Query Match	7.8%; Score 89.8; DB 8; Length 1681;			
Best Local Similarity	58.4%; Pred. No. 7.2e-11;			
Matches 157; Conservative	0; Mismatches 112; Indels 0; Gaps 0;			
Qy	672 tgaagaacatcagccatagagccaagaaggagatgggaacttctgaattgggaaga	731		
Db	988 TGAAGAGCAATATGCGAATAGGAAGATTAAGAGAGAGTCTCACTTGGGAGA	1047		
Qy	732 ttatcagaagatggaattcactcaatgctgcttctcagagcactgagtggtcaat	791		
Db	1048 TACAAAGAAAGATGCCATTACTTATAGAGTAATCCAAAGACATTAAAGAGTGCCTCAAT	1107		
Qy	792 cgtcaagactgtacatagaaaagctactcatatataataatcaagaatatgtattcc	851		
Db	1108 CTTATCTTTCACATTCACAGAGAGCTGCGAAGATGTGCAATACGAAGATATTGTATACC	1167		
Qy	852 aaaggaggaggaagtggttctcaatcttccagcagatcatctgttccctctcatga	911		
Db	1168 AAGAGGATGGAAGGTGTACCCCTATTCAGAAATATTCATGCAATGATGATATTTTTC	1227		
Qy	912 aaatcccttgaatttaattcccatgagat	940		
Db	1228 GGATCGGGGAAATTCGATCCGTCCAGAT	1256		

RESULT 14  
LOCUS OSJN00012/c 139969 bp DNA linear HTG 08-SEP-2001  
DEFINITION Oryza sativa chromosome 4 clone OSJNBa0089K21, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, in ordered pieces.  
ACCESSION AL606441  
VERSION AL606441.1 GI:1552679  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 139969)  
AUTHORS Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F.,  
Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,  
Chen, X. Y., Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W.,  
Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L.,  
Zhu, F. H., Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q.,  
Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Meng, Q. J.,  
Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y.,  
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K.,  
Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P.  
and Hong, G. F.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,  
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,  
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:  
bhan@ncgr.ac.cn  
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC  
clone: OSJNBa0089K21.  
COMMENT Web site: http://www.ncgr.ac.cn  
----- Summary Statistics  
Assembly program: phrap  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
source  
1. 139969  
/organism="Oryza sativa"  
/variety="Nipponbare"  
/sub-species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="4"  
/clone="OSJNBa0089K21"  
/clone\_1lb="COGI-OSJNBa"  
BASE COUNT 39133 a 31444 c 30948 g 38417 t 27 others  
ORIGIN  
Query Match 7.8%; Score 89.2; DB 2; Length 139969;  
Best Local Similarity 57.0%; Pred. No. 6.9e-11;  
Matches 163; Conservative 0; Mismatches 123; Indels 0; Gaps 0;  
OY 92 ggtatggaagaagtgttcagatcaatatatgtggtggaagaacagtgatcctatgtgacc 151  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 2033 ggtatgggggggttttaagctccatctgttcgacaccccccacatagtccttgacc 1974  
OY 152 aagaactcaacatgttcatactccaagaaggaaggaagtgttcacatcgattatccaa 211  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1973 aggaactgaacacattcattcattcagaatagagagagctgttttcagtcagctacc 1914  
OY 212 aagcagatgacatctcgcgcaaatatctcccttattatagcaccgcggaagaattaca 271  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1913 ggccaatcattgcattctgggcaagctccatgcttagtgcttccaggagagaccaca 1854  
OY 272 ggaactaaaaatgttatattatagcttcaatcacaagtcgaagcactgtcttc 331  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1853 agagctcaggaaccttgctcttagcagctagctccacaaagctcgaagccagctacc 1794

OY 332 ttcaactgcgcagagaacatctctatctcgatactaaatcagaa 377  
||| | ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1793 ttggccacatttgagaaagattgcactgcatttacttgggctatgcga 1748  
RESULT 15  
LOCUS OSJN00023/c 169845 bp DNA linear HTG 11-SEP-2001  
DEFINITION Oryza sativa chromosome 4 clone OSJNBa0016002, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, in ordered pieces.  
ACCESSION AL606588  
VERSION AL606588.1 GI:15994047  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 169845)  
AUTHORS Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F.,  
Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,  
Chen, X. Y., Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W.,  
Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L.,  
Zhu, F. H., Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q.,  
Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Meng, Q. J.,  
Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y.,  
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K.,  
Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P.  
and Hong, G. F.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,  
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,  
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:  
bhan@ncgr.ac.cn  
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC  
clone: OSJNBa0016002.  
COMMENT Web site: http://www.ncgr.ac.cn  
----- Summary Statistics  
Assembly program: phrap  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
source  
1. 169845  
/organism="Oryza sativa"  
/variety="Nipponbare"  
/sub-species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="4"  
/clone="OSJNBa0016002"  
/clone\_1lb="COGI-OSJNBa"  
BASE COUNT 47749 a 38566 c 37139 g 46370 t 21 others  
ORIGIN  
Query Match 7.8%; Score 89.2; DB 2; Length 169845;  
Best Local Similarity 57.0%; Pred. No. 6.8e-11;  
Matches 163; Conservative 0; Mismatches 123; Indels 0; Gaps 0;  
OY 92 ggtatggaagaagtgttcagatcaatatatgtggtggaagaacagtgatcctatgtgacc 151  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 165315 ggtatgggggggttttaagctccatctgttcgacaccccccacatagtccttgacc 165256  
OY 152 aagaactcaacatgttcatactccaagaaggaaggaagtgttcacatcgattatccaa 211  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 165255 aggaactgaacacattcattcattcagaatagagagagctgttttcagtcagctacc 165196  
OY 212 aagcagatgacatctcgcgcaaatatctcccttattatagcaccgcggaagaattaca 271  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 165195 ggccaatcattgcattctgggcaagctccatgcttagtgcttccaggagagaccaca 165136



THIS PAGE BLANK (USPTO)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 13.47 seconds

(without alignments)  
1096.060 Million cell updates/sec

Title: US-09-995-917A-1

Post-process score: 2027  
Sequence: 1 MGNPFIETISFFKPHRSDS.....YVEFKRGMLIEPTKFLIED 382

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SWISSPROT\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	655	32.3	472 1	C901_ARATH 042569 arabidopsis
2	638	31.5	524 1	C90C_ARATH 09m066 arabidopsis
3	568.5	28.0	464 1	CP85_LYCES 043147 lycopersico
4	432.5	21.3	490 1	C883_ARATH 023051 arabidopsis
5	390.5	19.3	519 1	C881_MAIZE 043246 zea mays (m
6	347.5	17.1	444 1	C120_SYNY3 059990 synechocyst
7	294.5	14.5	492 1	CP26_BRARE P79739 brachydanio
8	292.5	14.4	497 1	CP26_MOUSE 055127 mus musculu
9	280	13.8	492 1	CP26_XENLA 093323 xenopus lae
10	278.5	13.7	512 1	C26A_HUMAN 09nub4 gallus gall
11	275	13.6	482 1	CP26_CHICK 09nub4 gallus gall
12	267.5	13.2	487 1	CP26_HUMAN 043174 homo sapien
13	230.5	11.4	494 1	CPA7_HUMAN P20812 rattus norv
14	230.5	11.4	494 1	CPA7_HUMAN P20812 rattus norv
15	223.5	11.0	494 1	CPA4_MOUSE P15333 mus musculu
16	223	11.0	494 1	C912_ARATH 065790 arabidopsis
17	219.5	10.8	484 1	CPA5_MOUSE P20852 mus musculu
18	215.5	10.6	484 1	CPAC_HUMAN P11599 homo sapien
19	215.5	10.6	484 1	CPAB_RABIT 005356 oryctolagus
20	206.5	10.2	494 1	CPAB_RABIT 005356 oryctolagus
21	205.5	10.1	494 1	CPAB_MESAU P24454 mesocricetu
22	204.5	10.1	538 1	CP18_DROME 095078 drosophila
23	199	9.8	482 1	CPAC_MOUSE P56593 mus musculu
24	197	9.7	482 1	CPA1_RAT P17171 rattus norv
25	193	9.5	511 1	C7C4_ARATH 064655 arabidopsis
26	190.5	9.4	492 1	CPBC_RAT P33272 rattus norv
27	190	9.4	528 1	CP51_CANTR P14253 candida tiro
28	188.5	9.3	493 1	CPA9_MESAU P24455 mesocricetu
29	188.5	9.3	494 1	CPAD_HUMAN 016696 homo sapien
30	188	9.3	510 1	C755_EUSGR 096418 eustoma gra
31	187.5	9.3	494 1	CPBB_CANFA P24460 canis famill
32	186	9.2	528 1	CP51_CANAL P10613 candida alb
33	185.5	9.2	491 1	CPB4_RABIT P00178 oryctolagus

34	185.5	9.2	500 1	CPJ1_RABIT P52786 oryctolagus
35	185	9.1	492 1	CPA2_RAT P15149 rattus norv
36	185	9.1	504 1	CPB2_HELAM 027664 helioverpa
37	184.5	9.1	522 1	C822_SORYB 081972 glycine max
38	183.5	9.1	501 1	CP06_MOUSE 054750 mus musculu
39	182.5	9.0	490 1	CP07_RAT P05179 rattus norv
40	182.5	9.0	518 1	CP3R_ONCMY 042563 oncorhynch
41	182	9.0	508 1	CP48_RAT P24464 rattus norv
42	181.5	9.0	495 1	CAS3_DROME 09xyy0 drosophila
43	181.5	9.0	514 1	CP77_ICTRU 073853 icterulus p
44	181	8.9	500 1	C72V_ARATH 091196 arabidopsis
45	180.5	8.9	491 1	CPB5_RABIT P12789 oryctolagus

## ALIGNMENTS

```

RESULT 1
ID C901_ARATH STANDARD: PRT: 472 AA.
AC 042569:
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DE Cytochrome P450 90A1 (EC 1.14.-.-).
GN CYP90A1 OR CYP90 OR CPD OR AT5G05690 OR MJJ3.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=96200769; PubMed=8612270;
RA Szekeres M., Nemeth K., Koncz-Kalman Z., Mathur J., Kauschmann A.,
RA Altman T., Redei G.P., Nagy F., Schell J., Koncz G.;
RT "Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450,
RT controlling cell elongation and de-etiolation in Arabidopsis.";
RL Cell 85:171-182(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT pl clones.";
RL DNA Res. 4:215-230(1997).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X87367; CAA60793.1; -
CC EMBL: X87368; CAA60794.1; -
CC EMBL: AB005237; BAB0663.1; -
CC Interp: IPR001128; Cyt_P450.
CC Pfam: PF00067; P450_2.
CC DR PROSITE: PS00086; CYTOCHROME_P450; 1.
CC DR OXidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.
CC FT TRANSMEM 1 21 POTENTIAL.
CC FT BINDING 418 418 HEME (BY SIMILARITY).
CC SQ SEQUENCE 472 AA; 53785 MW; 41A73F46D64E343F CRC64;

```

Query Match 32.3%; Score 655; DB 1; Length 472;

Best Local Similarity 34.5%; Pred. No. 1.1e-37;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

0Y      1  MGNWPIGFIISFPRKHSRDSIGTFLQOAVSVYKGVFNKSNICGGAAVYSCDELMPFIION 60
Db      36  LGPLILGTFEOLIGAVKTEENPEPLDEWVARYSGVFMTHGEPTISADEPETERVYON 95
0Y      61  ESKLTETSDYPRKAMHDILCKYSLLATLAGEIHRKLNKVIISFINLRKSPDFLHCANESIS 120
Db      96  ESKLFECSYSPASICLLKHSLLMKGLSHLRKMSLTMSPFNSSITKDHMLDIDLRYRF 155
0Y      121  ILKSKNCREVEFHKRYKMFLLSVWNOLLSIKEDDPARLYVIDFLSYMKGFISLPIPL 180
Db      156  NLDSSWS--RYLIMEAKKIFTEFLTVKQMSFDGGEJMS- SLRREYLLVITRGFSLPL 212
0Y      181  PGTGTNAIKVRSNRNHHQNALI-----EDMNNALIREDFDSIISNED----- 222
Db      213  FSTYTRKAIQAR--RKVAEALTVVVMKRRKEEGBGAERKMKMLAALLAAADGSEDEIVD 270
0Y      225  -----EEHAAIRAKKGOGELLNMWEDYKME 249
Db      271  FLVALVAVAGYETTSIMTLAKVPLETPRLALAOIKEHEHETKIRAKKSDSYSIEMSDYSMP 330
0Y      250  FTQCVISEALBCGNIVKTHRKATHDIKFEKVEYLPKGMKVYPIPTAVHLDPSLHENPEEF 309
Db      331  FTQCVASENTELLRANIGGVFRAMTDEYKIKYLPKGMKVYSSPRAVHLDPNHFKDARTF 390
0Y      310  NPMKRTKI-----TAFGGGVRCPSGEGELKDLQATFLLHLVLVSYRM 350
Db      391  NPMKQNSVTTGPSNVEFTPGGSGPRICPGVEILARVALSVFLHSLVLTGFGSW 441

```

RESULT	2			
C90C_ARATH	STANDARD;	PRT;	524 AA.	
ID	C90C_ARATH			
AC	Q9M066; 023242;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Cytochrome P450 90C1 (EC 1.14...), (ROTUNDFOLIAR3).			
GN	ROT3 OR CYP90C1 OR At4G36380 OR C7A10.980 OR F23E13.220.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsids.			
NCBI_TaxID	3702;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98361880; PubMed=6694802;			
RA	Kim G.-T., Tsukaya H., Uchimida H.;			
RT	"The ROTUNDFOLIAR3 gene of Arabidopsis thaliana encodes a new member			
RT	of the cytochrome P-450 family that is required for the regulated			
RL	polar elongation of leaf cells.";			
LN	Genes Dev. 12:2381-2391(1998).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RX	MEDLINE=20083488; PubMed=10617198;			
RA	Mayer K.F.X., Schueller C., Wamutt R., Murphy G., Volckaert G.,			
RA	Poll T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,			
RA	Harris B., Anstorge W., Brandt P., Griwell L.A., Rieger M.,			
RA	Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,			
RA	Kreis M., Delseny M., Puginomech P., Watson M., Schmidheini T.,			
RA	Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,			
RA	Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,			
RA	Langham S.-A., McCallagh B., Bilham L., Robben J.,			
RA	Van der Schuren J., Grynoprez B., Chuang Y.-J., Vandenbusche F.,			
RA	Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,			
RA	Witzenecker T., Bothé G., Ramsperger U., Hilbert H., Braun M.,			
RA	Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,			
RA	Moollman P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P.,			
RA	Bernsteiner S., Hempel S., Feldpausch M., Lambrecht S., Van den Daele H.,			
RA	De Keyser A., Buyschaert C., Gtelen J., Villarroel R., De Clercq R.,			

RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,  
RA Petrell A., Rajandream M.A., Lyne M., Benes V., Reichmann S.,  
RA Borfova D., Bloeker H., Scharfe M., Grimm M., Loehnett T.-H.,  
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
RA Gabell C., Fuchs M., Fartmann B., Granderath K., Danner D., Herzl A.,  
RA Neumann S., Aitchison A., Vitale D., Lignotti R., Piravandi E.,  
RA Massenet C., Ouisley F., Claubaud G., Mendelein A., Peibler R.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Chedrol F., Cooke R., Berger C., Montfort A., Casacuberta E.,  
RA Gildoux T., Weber N., Vandenbol M., Baigues M., Terol J., Torres A.,  
RA Perez-Perez A., Purrelle B., Bent E., Jacques S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Belke C.,  
RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cortes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Hamon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Speth J., Ryan E., Andrews S., Gelsel C., Layman D.,  
RA Du H., All J., Bergloff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M., Martienssen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis*  
RT thaliana.";  
RL Nature 402:769-777(1999).  
CC -I- FUNCTION: MIGHT BE INVOLVED IN THE BIOSYNTHESIS OF STEROIDS  
CC IMPORTANT FOR THE POLAR ELONGATION OF CELLS DURING DEVELOPMENT.  
CC -I- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum  
CC (potential).  
CC -I- TISSUE SPECIFICITY: Ubiquitous.  
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AB008097; BAA37167.1; -  
DR EMBL: AL161589; CAB80304.1; ALT INT.  
DR EMBL: Z99708; CAB18650.1; ALT INT.  
DR EMBL: AL022141; CAA18139.1; ALT\_SEQ.  
DR InterPro: IPR001128; CYL\_P450.  
DR Pfam: PF00067; P450; 1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Transmembrane; Heme;  
KW Endoplasmic reticulum; Multigene family;  
FT TRANSMEM 4  
FT 24  
FT BINDING 463 463 HEME (BY SIMILARITY).  
SQ SEQUENCE 524 AA; 59389 MW; 550578908BDF272 C664;

Query Match	31.5%	Score 638	DB 1	Length 524
Best Local Similarity	32.2%	Pred. No. 1.8e-36		
Matches 141; Conservative	75	Mismatches 156	Indels 66	Gaps 7

[illegible]

OY 121 ILKSKNCREVEFEKVKMTLSVWVNOQLSIKPEDPARLYVLODFLSYMKGFISLPIPL 180  
 Db 194 TLASWACPLVHVODEIKKTFEILVKVLMSTSPGEDNMLKL -FEFEFIKGLICIPIKF 252  
 OY 181 PCTGTNAIKVRSNRNHNQNAIIEEDMNAI-----NEEDF- 215  
 Db 253 PCTRLYLKAKKERLIKVKVKVEEROYAMTTTSPANDVDVLLDGDGDEKOSQSPDFV 312  
 OY 216 ----LDSIISNED-----BEHAIRAKKD-GEILWMEYQ 246  
 Db 313 SKGIYEMMIPEEETPMTATLAVKFLSDNPVALAKLVEENEMKRRKLEGEETWTDYM 372  
 OY 247 KMEFTQCVISALRCGNIVTKVHRKATHDIKFEKVIIPKGMKVPPIETAVHLDPSLHENP 306  
 Db 373 SLSEFQVNIINELRMANIINGVMRKALDVEIKGYLIPKMCVLASFISVHMDIEDIDNP 432  
 OY 307 FEENPMRTKT-----TAFGGVRCVCGGELGKLOIAFLHHLVLSYRKIKSDE 356  
 Db 433 YQDFWMDRINGSANSSICFTPEGGGRLCPGLELSKLEISIFLHLVTRYSWTAEDE 492  
 OY 357 MPAHPVEFKGMLEI 374  
 Db 493 I-VSEPTVYMKRRRLPIRV 509

RESULT 3  
 CP85\_LYCES STANDARD: PRT: 464 AA.  
 ID CP85\_LYCES  
 AC O43147;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Cytochrome P450 85 (EC 1.14.-.-) (Dwarf protein).  
 GN CYP85 OR D.  
 OS Lycopersicon esculentum (Tomato).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RX MEDLINE=96266705; PubMed=8672892;  
 RA Bishop G.J., Harrison K., Jones J.G.D.;  
 RT "The tomato Dwarf gene isolated by heterologous transposon tagging  
 RT encodes the first member of a new cytochrome P450 family.";  
 RL Plant Cell 8:959-969(1996).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U54770; AAB17070.1; -;  
 DR Interpro: IPR001128; Cyt\_P450.  
 DR Pfam: PF00067; P450.1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450.1.  
 KM Oxidoreductase; Monooxygenase; Membrane; Heme.  
 FT BINDING 414 414 HEME (BY SIMILARITY).  
 SO SEQUENCE 464 AA; 53706 MW; D2B21AAB7B14E94 CMC64;

Query Match 28.0%; Score 568.5; DB 1; Length 464;  
 Best Local Similarity 30.2%; Pred. No. 8.6e-32;  
 Matches 130; Conservative 78; Mismatches 158; Indels 65; Gaps 9;  
 OY 1 MGWPIGTISTFRRHRSDSIG-TFLQORVSRVYKVFESNIGCGAAVYSCQELMFMFLQ 59  
 ||||| ||| | | :|::: ||||| |||: | :||| |||: |||

Db 39 MGWPLFGTEFEK-----LGPSEMKQORARYGSEFKSHILGCPITVSMSELNRYILV 92  
 OY 60 NEGRKFTSDYRKAMHDIIGKYSLLATGELIRKKNVSIISFNILTKSPDFLHCANLSI 119  
 Db 93 NEAKGLVGFQSMIDILGKCNIAAVNSAHKYMGCALLSLSPWIMDQLPKIDEFMR 152  
 OY 120 SILKSMKNCREVEFEKVKMTLSVWVNOQLSIKPEDPARLYVLODFLSYMKGFISLPIP 179  
 Db 153 SHLTWMDN-KVIDQEKNNKMAFLSSKQIAGIESTSLAQR-MSEFNVLVGTLSLPI 210  
 OY 180 LPTGTY-----TAAIKV-----RSNRNHNQNAIIEEDMNAI-----AIREDFLDST 219  
 Db 211 LPNTVHNGFQARKIIVMLTLIEBRASKEIQHDMGMYLNEBATRFKLLDDEMDILI 270  
 OY 220 IS-----NDEEHAIRAKKGGGELLMEDDYOKMFT 251  
 Db 271 ITILYSGEIVSTTSMMAVKTLHDHPKVLBELERKHMAREKKEDPDIDNDRSMFT 330  
 OY 252 QCVISEALRCGNIVTKVHRKATHDIKFEKVIIPKGMKVPPIETAVHLDPSLHENPFEENP 311  
 Db 331 RAVILETSRLATIVGVLRKTTQDMEINGIIPKGMRIYVYTRELNDPRLYPDPYSFP 390  
 OY 312 MRWTKT-----AFGGVRCVCGGELGKLOIAFLHHLVLSYRKIKSDEMIANPY 363  
 Db 391 WRWMDKSLHONSEFLVFGGTROCGKELGVAELSTFLHYFTVYKRWBEIGDGLMKRPR 450  
 OY 364 VEFKGMLEI 374  
 Db 451 VEAPNGLRIRV 461

RESULT 4  
 C883\_ARATH STANDARD: PRT: 490 AA.  
 ID C883\_ARATH  
 AC O23051;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cytochrome P450 88A3 (EC 1.14.-.-).  
 GN CYP88A3 OR AT1G05160 OR YIP8H12.23.  
 GN Arabidopsis thaliana (Mouse-ear cress).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eusteroids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RX MEDLINE=11130712;  
 RA White O., Alonso J., Altarfi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao O., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kuritz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Miltcher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its





```

Db 219 FSLPIDVPFSGLYRGLRARNFHSKIEENIRKKIODDNDNEBOKYKDALQLLIENSRSFD 278
OY 225 EEHAIRAKKGOGEL-----LNME DYK-----247
Db 279 EPPFSLOAMEAETELLEFGHETTASTATSLVFLGINTVEYOKVREEOEKVEMGYTPG 338
OY 248 -----MEFTQCVISEALRCGNIVTVHRRKATHDIFKEFYEPKMKVEPIPTAVH 297
Db 339 KGLSMELDLQAKYTGCVIETLRINPVPGGFRVALKTELANGYQIPKGMNVIYSICDH 398
OY 298 LDPSSHENPEFENPMKRWKT-----TAFGGVRCVCGGELGKLIQIAFFLHLVLVS 347
Db 399 DVADVPENKEEFQPERFEMKGLDGSRFNVYIPDGSGSRVCVGEKFAVLKIFLVELTQH 458
OY 348 YRRKIKSDDEM-----PIAHPIVEFRKGMLEIEPTKE 379
Db 459 CMMILSNGPPTMKGTPTTP-----VDNLPTKE 486

RESULT 8
CP26_MOUSE STANDARD; PRT; 497 AA.
ID CP26_MOUSE 05127; Q9R1F4;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid-metabolizing
  cytochrome) (P450RA1) (Retinoic acid 4-hydroxylase).
GN CYP26A1 OR CYP26 OR P450RA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=9792446; PubMed=9250660;
RA Fujii H., Sato T., Kaneko S., Gotoh O., Fujii-Kuriyama Y., Osawa K.,
  Kato S., Hamada H.;
RT "Metabolic inactivation of retinoic acid by a novel P450
  differentially expressed in developing mouse embryos.";
RL EMBO J. 16:4163-4173(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9811212; PubMed=9442090;
RA Abu-Abed S.S., Beckett B.R., Chiba H., Chithalen J.V., Jones G.,
  Metzger D., Chambon P., Petkovich M.;
RT "Mouse P450RA1 (CYP26) expression and retinoic acid-inducible retinoic
  acid metabolism in F9 cells are regulated by retinoic acid receptor
  gamma and retinoid X receptor alpha.";
RL J. Biol. Chem. 273:2409-2415(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Tooth;
RX MEDLINE=20515038; PubMed=11063033;
RA Paine C.T., Paine M.L., Shead M.L.;
RT "Identification of tufelrin- and amelogenin-interacting proteins using
  the yeast two-hybrid system.";
RL Connect. Tissue Res. 38:257-267(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays a key role in retinoic acid metabolism. Acts on
  retinoids, including all-trans-retinoic acid (RA) and its
  stereoisomer 9-cis-RA. Capable of both 4-hydroxylation and 18-
  hydroxylation. Responsible for generation of several hydroxylated
  forms of RA, including 4-OH-RA, 4-oxo-RA, AND 18-OH-RA.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDICATION: BY RETINOIC ACID.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: Y12657; CA973206.1; -.
DR EMBL: AF115769; AAD17217.1; -.
DR EMBL: BC012673; AAH12673.1; -.
DR MGD: MGI:1096359; Cyp26.
DR InterPro: IPR001128; CYP_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
  Endoplasmic reticulum.
FT BINDING 442 442 HEME (POTENTIAL).
FT CONFLICT 9 9 S -> T (IN REF. 3).
FT CONFLICT 154 154 L -> P (IN REF. 4).
FT CONFLICT 356 356 I -> T (IN REF. 4).
FT CONFLICT 492 492 Y -> H (IN REF. 4).
SQ SEQUENCE 497 AA; 56177 MW; 3B07D7C29134471 CRC64;

Query Match 14.4%; Score 292.5; DB 1; Length 497;
Best Local Similarity 23.2%; Pred. No. 5,7e-13;
Matches 106; Conservative 61; Mismatches 187; Indels 103; Gaps 10;

OY 1 MGNPIGETISFEKPHRSISIGTFLQQRVSRGYKFKSNICGKAVVSCDQLNMFILON 60
  ||:||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 49 MGPFEGETLQWVLORRK-----FLQMKRRKYGFYIKTHLFRPVRVAGADVNRILLG 103
OY 61 ECKLFTSDYPRKAMHDLKGYSLILATGEIHRKIKANYIFINLTKSKPFLMCAENLTS 120
  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 104 EHRLLSVHWPASVRTLIGAGCLSNLHDSHKQKKVIMQAFSEALQCYLVIAEVS-S 162
OY 121 ILSKWKNC--REVEFHKEMKFLSVNQLSIRP-----EDPARLVLDPLSYMG 172
  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 163 CLEQWLSCGERGLLYPEVKRLMFRIMKILLGCEGPRAGGDEQL--VEAFEMFTN 220
OY 173 FLSLPIDPFGTYNAIKVRSNRNIHQNAIIEEDMNAIE-----EDFLDSITSN 222
  ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 221 LFLPIDVPFSGLYRGVKA---NLIHRIEENIRAKIRRLQATEPDGCKALQLLIH 277
OY 223 -----EDEHAIRA 232
Db 278 SWERGERLDMQALKOSSTELLEFGHETTASATSLITYGLIPHYLQKYREELKSKGLIC 337
OY 233 KRGDGELLNMEDEYOKMEFTQCVISEALRCGNIVTVHRRKATHDIFKEFYEPKMKVPEP 292
  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 338 KSNQDKLDMETLEQLKIGCVIKETLRINPVPGGFRVALKTELANGYQIPKGMNVIYS 397
OY 293 FTAVHLDPSSLHENPEFENPMKRWKT-----KTAFGGVRCVCGGELGKLIQIAFFLH 342
  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 398 ICDTHVADIFETNKEEFNDRFIVPHPEDASRSPFIPFGGGLSCVGEKFAVLKIFTV 457
OY 343 HVLVSYRRKIKS-----DEMPIAHPIVE 365
Db 458 ELARHCDWOLLNGPPTMKTSPIVYVDNLPARFTYQ 494

RESULT 9
CP26_XENLA STANDARD; PRT; 492 AA.
ID CP26_XENLA 09323;
AC 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DR 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid degrading enzyme
  CYP26) (XCP26) (Retinoic acid converting enzyme) (RACE).
GN CYP26A1 OR CYP26.

```

OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_Taxid=8355;  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99077803; PubMed=9857192;  
RA Hollemann T., Chen Y., Grunz H., Pleier T.;  
RT "Regionalized metabolic activity establishes boundaries of retinolic acid signaling";  
RL EMO J. 17:7361-7372(1998).  
CC  
CC -1- FUNCTION: Plays a key role in retinolic acid metabolism (By similarity).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By similarity).  
CC -1- INDUCTION: By retinolic acid.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/announce/or\\_send\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_email_to_license@isb-sib.ch)).  
CC  
DR EMBL: AF057566; AAC25158.1; -  
DR InterPro: IPR001128; Cyt\_P450.  
DR Pfam: PF00067; P450.1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450.1.  
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome; Endoplasmic reticulum.  
FT BINDING 437 437 HEME (POTENTIAL).  
FT SEQUENCE 492 AA; 55459 MW; D1D4BB7651BF2D3E CRC64;  
Query Match 13.8%; Score 280; DB 1; Length 492;  
Best Local Similarity 24.1%; Pred. No. 4e-12;  
Matches 104; Conservative 54; Mismatches 176; Indels 98; Gaps 12;  
OY 1 MGPPIGTISFFKPHRSISGTFLOQVRSRGKVKSGICGKAVSCDOLNFIION 60  
DB 49 MGLPFEGCELOWLQRR-----RFLQVRSQYGRITVTHLFGSPVRYTGAVNRQILMG 103  
OY 61 BECKLTSDPKAMHDLGKYSLLATGEIHRKLVNI-----SPINLTGSKPFIICA 114  
DB 104 EKKLVSVHMPASVRIILGACLSNLHDEHKYTKVINAQSFREALANTVPMEEVRCSS 163  
OY 115 ENLISILSKMKN---CREVEFHKVKMTLSVMNQLSIRKEDPARL-----VYLQD 165  
DB 164 VNL-----WLGSPC--VLVYPAIKRMFRIAMRLILGCG--DPQRMDRQOEETLLEA 211  
OY 166 FLSTYKGFSLPIPLPGGYTNAIVR----- 192  
DB 212 FEEMSRNLFSLPIDVPFSGLYRGLARNLHAQIEENIKELQREDEHCKDALQILLDY 271  
OY 193 SNRN---IHONAIIDMNNAI-----REDFLDSIISNEDEHNAI 230  
DB 272 SRNRNEPILNQLAKESATELLFGGHTTASATSLTSPALAKDVEK--VRKELEQGL 330  
OY 231 RAKKGDGLLNEDYOKMEFTQCVISEALRCGNIVKTVHRKATHDIKFEKVIYPRGMKVF 290  
DB 331 STKPEKKELSTEVLQQLKTSYCIKETRLRSPVAGGFRVALKTFVLNGYDIPKGMVYI 390  
OY 291 PLFTAVHLDPSLHNPFEENPKRMKT-----TAFGGGVRCPCGGELGKQIDAF 340  
DB 391 YSIADTGHGADLPPTDKFNPRFLPLPRDSRSGGFIFPGGVCICGKEFAKILKVF 450  
OY 341 LHHVLVYRWKI 352  
DB 451 VVELRCNDMEL 462

RESULT 10  
C26A\_HUMAN  
ID C26A\_HUMAN STANDARD; PRT; 512 AA.  
AC Q9NR63; Q9NP41; -  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Cytochrome P450 26A2 (EC 1.14.-.-) (P450RA1-2) (Retinolic-acid metabolizing cytochrome).  
DE  
GN CYP26A2 OR P450RA12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Retina;  
RX MEDLINE=20300913; PubMed=10823918;  
RA White J.A., Ramshaw H., Taimi M., Stangle W., Zhang A., Everingham S., Creighton S., Tam S.-P., Jones G., Petkovich M.;  
RT "Identification of the human cytochrome P450, P450RA1-2, which is predominantly expressed in the adult cerebellum and is responsible for all-trans-retinolic acid metabolism";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:6403-6408(2000).  
RN  
RP SEQUENCE FROM N.A.  
RA Bourne S., Bauer C., Pape K., Jones T.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Plays a key role in retinolic acid metabolism. Involved in the specific inactivation of all-trans-RA. Responsible for generation of several hydroxylated forms of RA, including 4-OH-RA, 4-oxo-RA, and 18-OH-RA.  
CC -1- ENZYME REGULATION: Has a preferred activity towards the following substrates: all-trans-RA > 9-cis-RA > 13-cis-RA.  
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By similarity).  
CC -1- TISSUE SPECIFICITY: Highly expressed in brain, particularly in the cerebellum and pons.  
CC -1- INDUCTION: By retinolic acids (RA).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/announce/or\\_send\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_email_to_license@isb-sib.ch)).  
CC  
DR EMBL: AF252297; AAF76003.1; -  
DR EMBL: AC007002; AAF65576.1; -  
DR MIN: 605207; -  
DR InterPro: IPR001128; Cyt\_P450.  
DR Pfam: PF00067; P450.1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450.1.  
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome; Endoplasmic reticulum.  
FT BINDING 441 441 HEME (POTENTIAL).  
FT CONFLICT 68 68 Q -> OKCTLRVWMLQ (IN REF. 2).  
FT SEQUENCE 512 AA; 57512 MW; A06D1D9944E6726F CRC64;  
Query Match 13.7%; Score 278.5; DB 1; Length 512;  
Best Local Similarity 23.0%; Pred. No. 5.3e-12;  
Matches 105; Conservative 69; Mismatches 161; Indels 121; Gaps 15;  
OY 1 MGPPIGTISFFKPHRSISGTFLOQVRSRGKVKSGICGKAVSCDOLNFIION 60  
DB 54 MGPPIGTG-----GHMLDGGSGFSSRRKRYGNVFKTHLGRPLIRVGAENVRIILMG 108

	RESULT	11
CC	CP26.CHICK	
ID	CP26.CHICK	STANDARD:
AC	Q9P0B4;	PRT; 492 AA.
DT	01-MAR-2002 (Rel. 41, Created)	
DT	01-MAR-2002 (Rel. 41, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid degrading enzyme CYP26).	
GN	CYP26A1 OR CYP26.	
OS	Gallus gallus (Chicken).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
OC	Gallus.	
OX	NCBI_TaxID=9031;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryonic spinal cord;	
RX	MEDLINE=20057772; PubMed=10588879;	
RA	Syndell E.C., Thaller C., Sockanathan S., Petkovich M., Jessell T.M.	
RA	Etchebe G.;	
RT	"Complementary domains of retinoic acid production and degradation in the early chick embryo.";	
RL	Dev. Biol. 216:282-296(1999).	
RN	[2]	
RP	SEQUENCE OF 368-474 FROM N.A.	
RA	Martinez-Ceballos F., Burdzel C.A.;	
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.	
-1-	FUNCTION: Plays a key role in retinoic acid metabolism. Appears to be involved in the specific inactivation of all-trans-RA. Responsible for generation of several hydroxylated forms of RA, including 4-OH-RA, 4-oxo-RA, and 18-OH-RA.	
-1-	SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.	
-1-	DEVELOPMENTAL STAGE: Expressed at stage 4 in the ectoderm, stage 5-7 in the nascent notochord and at stage 7 its expression decreases in the anterior part of the embryo. From stage 7-10 its expression is restricted to the dorsal folds of the neural tube and to rhombomere 2. At stage 10, it is expressed in the lateral plate endoderm and in the tail bud and by stage 11/12 it disappears in the neural tube, followed by a confined expression at stage 12 to dorsal neural tube and at stage 15 an increasing expression in the ectoderm.	
-1-	INDUCTION: By retinoic acid.	
-1-	SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.	

Query Match	13.6%;	Score 275;	DB 1;	Length 492;
Best Local Similarity	22.6%;	Pred. No. 8.8e-12;		
Matches 106;	Conservative 64;	Mismatches 178;	Indels 122;	Gaps 13;

```

QY      1 MGWPEIGETISFFPKPHRSDSICGTQOVSRYGKFKYKSGAAVYSCDELNMFIION 60
Db      49 MGLPFGGTTLMVLORR-----FLQMRKRYKGYIKNHLGRTVRYVMGAEVNRHILLG 103
QY      61 ECKLETSYPRAMHDILGKYSILLATGCIHRKLKAVIISFINLTKSKPDELH----CAEN 116
Db      104 EHRUVSYWMPASVRLTIGSGCLSNLHNGQHRKRVIMQAF-----SRDALHYVYVIOEE 159
QY      117 LSIISLKSMMK-----CREVEPKVEYKMFGLSMVMQOLISIKPEDPA---RLVVLDDFLSYM 170
Db      160 VS-ACLAOWLQAGRCFLV--TPYVKRLMFRAMRILLGFGFRQASPDGEQQLVAEFEMTI 216
QY      171 KGFISLPILDEGTGYNAIKVRS--NRNINON----- 200
Db      217 RNLSFLPIDVPFSGHYRLRARNIITHAKIEENIRAKMARKEPEGYKDALQLMEHTQGN 276
QY      201 -----ATIEDNMNAIREDPFDLSIISNDE 225
Db      277 GEOLNMOELKESATELFFGHEHTTASAASTSLIAELGLNHDLQVKARELQJLKGSPNQ 336
QY      226 EHAHAIRAKKGDELNMEDYQKMEFTQCVISEALRCGNIVTVYRHKATHIDIKFEYUVPK 285
Db      337 E-----KOLNMFEQLKTYGIVIKETLRLSPRYGGRIRIAKLTLELNGYOIRK 385
QY      286 GWKVPFIFTAVHLDPSCLEHNPFEFNPMKRYKT-----AFGGGVRYVCPGGELGKL 335
Db      386 GWNYIYSICDTHDAVDLFTDKDEFNPDRFMSPSDESSRSFGRFPGGGGLRSVCYGEKFAKV 445
QY      336 QIAFFLHLVLVTSRWKIKSDM-----PIADRYVEFKRMGLLELEPRTFL 380
Db      446 LKLFYVELANSCDWOLNGLNPPTKMGTYVP-----VNDUPAFETI 486

```

RESULT	12
CP26_HUMAN	
ID	CP26_HUMAN
AC	043174;
DT	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid-metabolizing cytochrome) (P450RAI) (hp450RAI) (Retinoic acid 4-hydroxylase) .
GN	CYP26A1 OR CYP26.
OS	Homo sapiens (Human) .
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



```

OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97373542; PubMed=9228017;
RA White J.A., Beckett-Jones B., Guo Y.-D., Dllworth F.J., Bonasoro J.,
RA Jones G., Petkovich M.,
RT "CDNA cloning of human retinoid acid-metabolizing enzyme (hP450RA1)
RT identifies a novel family of cytochromes P450."
RL J. Biol. Chem. 272:18538-18541(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98380037; PubMed=9716180;
RA Somerville E., van den Brink C.E., van der Leede B.M., Schukkes R.K.,
RA Petkovich M., van der Burg B., van der Saag P.T.,
RT "Human retinoid acid (RA) 4-hydroxylase (CYP26) is highly specific for
RT all-trans-RA and can be induced through RA receptors in human breast
RT and colon carcinoma cells."
RL Cell Growth Differ. 9:629-637(1998).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=99045433; PubMed=9826557;
RA Trofimova-Griffin M.E., Juchau M.R.;
RT "Expression of cytochrome P450RA1 (CYP26) in human fetal hepatic and
RT cephalic tissues."
RL Biochem. Biophys. Res. Commun. 252:487-491(1998).
CC -1- FUNCTION: Plays a key role in retinoid acid metabolism. Acts on
CC retinoids, including all-trans-retinoic acid (RA) and its
CC stereoisomer 9-cis-RA. Capable of both 4-hydroxylation and 18-
CC hydroxylation. Responsible for generation of several hydroxylated
CC forms of RA, including 4-OH-RA, 4-oxo-RA, AND 18-OH-RA.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: Highest levels in adult liver, heart,
CC pituitary gland, adrenal gland, placenta and regions of the brain.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF005418; AAB88881.1; -.
DR HSSP: P33006; 1CPT.
DR MIM: 602239; -.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PFO0067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
KW Endoplasmic reticulum.
FT BINDING 442 442 HEME (POTENTIAL).
SQ SEQUENCE 497 AA; 56162 MW; EABBB84B24B2EAB3 CRC64;

```

```

Query Match 13 28; Score 267.5; DB 1; Length 497;
Best Local Similarity 23.28; Pred. No. 2.9e-11;
Matches 100; Conservative 63; Mismatches 177; Indels 91; Gaps 13;

```

```

OY 1 MGFPIGETTISFFKPHRSDSIGTFLQORVSRGKVFKNICGKAVVSCDOELMFTLION 60
OY 49 MGFFPGETLQVLOVRK-----FLQKKRRKRYGTYTHLEGRTVYMGADNVRILLG 103
OY 61 EGGLETSDYPRAMHDILGKYSLLATGELIRKLNKVIISFTNLTKSRPDLIHC-----AE 115
OY 104 DDRIVSVMPASVRYTIIISGCLSNLHDSHQKRVIMRAPS-----REALECVPYITE 158
OY 116 NLSTISLTKMKNC--REVEFHKEVKMFTLSVMNOLLSTIKP-----EDPARLYVQDPLSY 169
OY 159 EVG-SLEQWLSGCGERGLLVPEVKRLMFRITAMRILLGCEPQLAGDDSEQLVEAFEE 217

```

```

OY 170 MKGISLPILPGTGYNAIKVRS--NRNIHON-----AILE- 204
OY 218 TRNLFSLPDPFSSLYGKMRANLIRHRIENIAKICGLRASRAGCKDALQDLLEH 277
OY 205 -----DMNNAIRREE-----DFLSIIS-----NEDESHAIR 231
OY 278 SWERGERLDM-QALKOSSTELFGGHETTASAASTLITLYGLYPHYLVKQVRELSKGL 336
OY 233 AKKGSGELLNMDYQKMETOCVISEALRCGNIVTVIRKANHDIKFEYVIPKGMKVP 291
OY 337 CKSNODNKLDMELILQKIGCVIKETLRNPVPGGFWALKTEFLNGYQIPKGMNVIY 396
OY 292 IFTAVHLPDSLHENPFENPMRMT-----KTTAFGGGVRVCPGGELGKQIAFFL 341
OY 397 SICDTHDAVEITNNNEFFRDPFMLRPHEDASRFSPFPGGELRSGVGEFAKILKIY 456
OY 342 HHVLVSYRWKI 352
OY 457 VELARHCDWQL 467

```

RESULT 13

```

ID CP43_RAT STANDARD; PRT; 494 AA.
AC P20812;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome P450 2A3 (EC 1.14.14.1) (CYP1A3) (Coumarin 7-hydroxylase).
GN CYP2A3 OR CYP2A-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;
RX MEDLINE=90356430; PubMed=2388852;
RA Ueno T., Gonzalez F.;
RT "Complete sequence of the rat CYP2A3 gene specifically transcribed in
RT lung."
RL Nucleic Acids Res. 18:4623-4624(1990).
RN [2]
RP SEQUENCE OF 8-494 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=89323084; PubMed=2751996;
RA Kimura S., Kozak C.A., Gonzalez F.J.;
RT "Identification of a novel P450 expressed in rat lung: cDNA cloning
RT and sequence, chromosome mapping, and induction by
RT 3-methylcholanthrene."
RL Biochemistry 28:3798-3803(1989).
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC ACIDS, AND XENOBIOTICS.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: LUNG.
CC -1- INDUCTION: BY 3-METHYLCOLANTHRENE (3MC).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M33190; AAA41022.1; -.
DR EMBL: J02852; AAA88511.1; -.

```

DR PIR; A32030; A32030.  
 DR PIR; S15056; S15056.  
 DR HSSP; P00179; 1DT6.  
 DR InterPro: IPR001128; Cyt\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT BINDING 439 439 HEME.  
 FT CONFLICT 109 109 W->G (IN REF. 2).  
 FT CONFLICT 335 335 I->N (IN REF. 2).  
 SQ SEQUENCE 494 AA; 56510 MW; F71855CBE602672F CRC64;

Query Match 11.4%; Score 230.5; DB 1; Length 494;  
 Best Local Similarity 22.4%; Pred. No. 9.4e-09;  
 Matches 102; Conservative 67; Mismatches 147; Indels 139; Gaps 14;

QY 4 PFIGETISFFKPHRSDISGTFLOQRVSRGKVFKEKSNICGKAVSCDQELNMFILQNEGK 63  
 DB 41 PFIGNYIQL-----NTEKWSYSLMKISQRYGVFTIHLGPRRVVVLGGDAVKEALVDQAE 96  
 QY 64 LFTSDYKAMHDILGK-YSLLATGELHKKIKNVIIISFINLTSKPDPLHCAENLSTISIL 122  
 DB 97 EFSGRGEQATFDMVFKGYGVAFFSSGERAKQLR-----RFSIATL 135  
 QY 123 KSW---KNCREVEFEKFEV-----KMFTLSVMNQLLSTIKPEPARLYVL 163  
 DB 136 RDGVGKRGIEERIQEAGFLIESFKTNGALDPTTYSLSVNSVYFVGDRPED 195  
 QY 164 QDLSTYK-----GFISLPPLPG-----GYTNAIKVRSNRN 196  
 DB 196 KEFTSLIRMMIGSFQFATSTGOLYEMFSSVMKHLPGQQAFAKGLGLEDFT---TKK 251  
 QY 197 IHONAIIEDNNNAIREDDEFDSIISNDEE----- 226  
 DB 252 VEONQRTLDNS---PDFIDISFLIRMLEKKKNPTEFYKKNLVLTLLNFFAGTETVST 308  
 QY 227 -----HAIIRAK-----GDGELLNMEYQKMEFTQCVISALRCGNIVTK 267  
 DB 309 TLRYGFLLMKHPDIEKVAHEIDRYIGNRQAKYEDRMKMPTEAVTIELQFADNIPM 368  
 QY 268 -VIRKATHDIKEKEYIYIPKWKVPIFTAVHLDPDSLHENPEFNPW-----RWTKTTA 319  
 DB 369 GLARVYKDKTKREFFLPKGTEVFPMIGSVLKDPKFFSNPNDEPNPKHFLDCKQFKKSDA 428  
 QY 320 ---FGGGRVCPGSGELGKLIQIAFFLHHLVLSYRK 351  
 DB 429 FVPSISGKRYCPSGEGLARMLFLFLINIMONFCFK 463

RESULT 14  
 ID CPAT\_HUMAN STANDARD; PRT; 494 AA.  
 AC P20853; O13121;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Cytochrome P450 2A7 (EC 1.14.14.1) (CYP1A7) (P450-IIA4).  
 GN CYP2A7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90212623; PubMed=2322567;  
 RA Yamano S., Tatsuno J., Gonzalez F.J.;  
 RT "The CYP2A7 gene product catalyzes coumarin 7-hydroxylation in human  
 liver microsomes.";  
 RL Biochemistry 29:1322-1329(1990).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95397851; PubMed=7668294;  
 RA Fernandez-Salguero P., Hoffman S.M., Cholesterol S., Mohrenweiser H.,  
 RA Raunio H., Rautio A., Pelkonen O., Huang J.D., Evans W.E., Idle J.R.;  
 RT "A genetic polymorphism in coumarin 7-hydroxylation: sequence of the  
 RT human CYP2A genes and identification of variant CYP2A6 alleles";  
 RL Am. J. Hum. Genet. 57:651-660(1995).  
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
 CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
 CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
 CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
 CC ACIDS, AND XENOGENS.  
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER  
 CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,  
 CC AND CARCINOGENS.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M33317; AAA52138.1; -;  
 DR EMBL; U22029; AAB40520.1; -;  
 DR PIR; C34271; C34271.  
 DR HSSP; P00179; 1DT6.  
 DR MIM; 123960; -;  
 DR InterPro; IPR001128; Cyt\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT BINDING 439 439 HEME (BY SIMILARITY).  
 FT CONFLICT 108 108 D->E (IN REF. 2).  
 FT CONFLICT 163 163 S->T (IN REF. 2).  
 FT CONFLICT 226 226 V->L (IN REF. 2).  
 FT CONFLICT 242 242 Q->L (IN REF. 2).  
 FT CONFLICT 402 403 PS->LR (IN REF. 2).  
 FT CONFLICT 409 409 Q->R (IN REF. 2).  
 FT CONFLICT 418 419 DD->GE (IN REF. 2).  
 FT CONFLICT 426 426 S->R (IN REF. 2).  
 FT CONFLICT 435 435 G->R (IN REF. 2).  
 FT CONFLICT 438 438 Y->N (IN REF. 2).  
 FT CONFLICT 475 475 S->P (IN REF. 2).  
 FT CONFLICT 479 479 G->V (IN REF. 2).  
 SQ SEQUENCE 494 AA; 56408 MW; 4094386AA07F9016 CRC64;

Query Match 11.4%; Score 230.5; DB 1; Length 494;  
 Best Local Similarity 23.7%; Pred. No. 9.4e-09;  
 Matches 108; Conservative 62; Mismatches 157; Indels 129; Gaps 19;

QY 4 PFIGETISFFKPHRSDISGTFLOQRVSRGKVFKEKSNICGKAVSCDQELNMFILQNEGK 63  
 DB 41 PFIGNYIQLNTEHICDSIMKFS-----CYGVFTIHLGPRRVVVLGGDAVKEALVDQAE 96  
 QY 64 LFTSDYKAMHDILGK-YSLLATGELHKKIKNVIIISFINLTSKPDPLHCAENLSTISIL 122  
 DB 97 EFSGRGEQATFDMVFKGYGVAFFSSGERAKQLRFAFATLRDGVGKRGIEERIOESGFL 156  
 QY 101 INTKSK-----PDFL--HCAENLSTILSKMKKNREVEFEKVMF--TTSVMNQLL 150  
 DB 157 IEAIRSHCANIDPTFFLSRTYSNVYSSIVFGDR-----FYEDEKFSLLSMIG-IF 209  
 QY 151 SIKPEDPARLYVLQDLSTYMKGFISLPI-----LPQGYTNAIKVRSN-RNIHONAI 202

Db 210 QFTSTGTGQLEEM--ESSVVK---HLPGPQQAQAKLLQGLEDFLAKKVEHNORTLDPS- 263  
QY 203 IEDMNNAIREDDELDSIISNEDEE----- 226  
Db 264 -----PQDFIDSLFIMQEEKKNPTEFLKLNLMSTLNFINGTETVSTTLRYGF 314  
QY 227 -----HAATIAKK-----GGELLNMDYOKMEFTQCVISEALRCGNIV-KTVHKA 272  
Db 315 LLLMKHPEVAKVHEEDIDRVIGKNNROKPFEDRTKMPYMEAVIHQIFGDIPIKSLARV 374  
QY 273 THDIKFEYVIRPKMKVFPFTAVHIDPSLHNPFEENPM-----RMTKTA---FGG 322  
Db 375 KQTKRFDFLPKGTETVPEMLGSLVRDPSFSPNQDFNPDHFLDKQGFKKSADAFVPEST 434  
QY 323 GVAVCPGSELGKQIAFFLHHLVLSYRKIKSDMP 358  
Db 435 GKRYCFCGGLARMEEL--ELEFTYVMQNFRLKSSOSP 468

RESULT 15  
CP4A\_MOUSE  
ID CP4A\_MOUSE STANDARD: PRT: 494 AA.  
AC P15392;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cytochrome P450 2A4 (EC 1.14.14.1) (CYP11A4) (Testosterone 15-alpha-hydroxylase) (P450-15-alpha) (P450-11A3.1).  
CN CYP2A4 OR CYP2A-4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA MEDLINE=89197954; PubMed=2703500;  
RA Lindberg R., Burkhart B., Ichikawa T., Negishi M.;  
RT "The structure and characterization of type I P-450(15) alpha gene as major steroid 15 alpha-hydroxylase and its comparison with type II P-450(15) alpha gene.";  
RT J. Biol. Chem. 264:6465-6471(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88153731; PubMed=3346244;  
RA Squitres E.J., Negishi M.;  
RT "Reciprocal regulation of sex-dependent expression of testosterone 15 alpha-hydroxylase (P-450(15) alpha) in liver and kidney of male mice by androgen. Evidence for a single gene.";  
RT J. Biol. Chem. 263:4166-4171(1988).  
RN [3]  
RP MUTAGENESIS.  
RX MEDLINE=89281737; PubMed=2733794;  
RA Lindberg R., Negishi M.;  
RT "Alteration of mouse cytochrome P450och substrate specificity by mutation of a single amino-acid residue.";  
RT Nature 339:632-634(1989).  
RN [4]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=99421934; PubMed=10490589;  
RA Lavery D.J., Lopez-Molina L., Margueron R., Fleury-Olela F., Conquet F., Schibler U., Bonfils C.;  
RT "Circadian expression of the steroid 15 alpha-hydroxylase (Cyp2a4) and coumarin 7-hydroxylase (Cyp2a5) genes in mouse liver is regulated by the PAR leucine zipper transcription factor DBP.";  
RT Mol. Cell. Biol. 19:6488-6499(1999).  
CC -1- FUNCTION: HIGHLY ACTIVE IN THE 15-ALPHA-HYDROXYLATION OF TESTOSTERONE. ALSO ACTIVE IN THE 15-ALPHA-HYDROXYLATION OF PROGESTERONE AND ANDROSTENEDIONE. LITTLE OR NO ACTIVITY ON CORTICOSTERONE, PREGNENOLONE, DEHYDROPIANDROSTERONE, ESTRADIOL OR ESTRIDOL.  
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

CC oxidized flavoprotein + H(2)O.  
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC TISSUE SPECIFICITY: KIDNEY AND LUNG. EXPRESSED IN LIVER, WITH A  
CC STRONG CIRCADIAN RHYTHMICITY. CIRCADIAN EXPRESSION IS REGULATED BY DBP.  
CC -1- MISCELLANEOUS: THERE ARE ONLY 11 DIFFERENCES BETWEEN THE SEQUENCE OF TESTOSTERONE 15-ALPHA-HYDROXYLASE AND THAT OF COUMARIN 7-HYDROXYLASE. BY SITE-DIRECTED MUTAGENESIS IT HAS BEEN SHOWN THAT MODIFICATION OF POSITION 209 IS SUFFICIENT TO CONVERT THE SPECIFICITY OF THE TWO FORMS OF THE ENZYME.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: M26208; AAA37797.1; -  
DR EMBL: M25146; AAA37797.1; JOINED.  
DR EMBL: M25147; AAA37797.1; JOINED.  
DR EMBL: M26202; AAA37797.1; JOINED.  
DR EMBL: M26203; AAA37797.1; JOINED.  
DR EMBL: M26205; AAA37797.1; JOINED.  
DR EMBL: M26206; AAA37797.1; JOINED.  
DR EMBL: M26207; AAA37797.1; JOINED.  
DR EMBL: J03549; AAA40426.1; -  
DR EMBL: M19319; AAA40429.1; -  
DR PIR: A33531; A33531.  
DR HSSP: P00179; 1D76.  
DR MGI: 88596; Cyp2a4.  
DR InterPro: IPR001128; CyL\_P450.  
DR Pfam: PF00067; P450; 1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
DR KX Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 439 439 HEME.  
FT CONFLICT 296 306 L -> Q (1N REF. 2).  
FT CONFLICT 306 306 V -> G (1N REF. 2).  
SQ SEQUENCE 494 AA; 56594 MW; 69724BDE3195D75C CRC64;

Query Match 11.0%; Score 223.5; DB 1; Length 494;  
Best Local Similarity 22.3%; Pred. No. 2.8e-08;  
Matches 100; Conservative 69; Mismatches 154; Indels 125; Gaps 14;

QY 4 PFIGETISFEKPHRSDISGTFLOQVRSGYKVPKSNICGKAVSCDQELNMFILQNEGK 63  
Db 41 PFVGN-----FLQNTQEQMNSLMKISQRYGPTTYILGSRIRYVLCQGEVAKVALDOAE 96  
QY 64 LFTSDYPKAMHDLGK-YSLLATGSEIHRKLNVIISFI----- 101  
Db 97 EESGREGQATFWLFFKGYGIAFSSGERAKQRLRSFATLRDFGVGRKIEERIQEAGFL 156  
QY 102 --NLTKSPKDFL-----HCAENLSITLSKSKNCRREVEFKHVKMFLSVAVNQLLSI 152  
Db 157 IDSFRKTNCAFDPTPTLYSLRTVSNVISTVFGDR-----FDVEDEF--LSLRMLGLSL 209  
QY 153 KPEDPARLVLDPLFSYMGFISLPRLPG-----TGYNNAKVSNNRIHONAI 203  
Db 210 QFVATSMGVVEFSSVMK-----HLPGPQQAQAKELQGLEDF-----TKKVEHNQRT 258  
QY 204 EDMNNAIREDDELDSIISNEDEE----- 226  
Db 259 LDPSN---PRDFIDSLFIMLEKKNPTEFYMKNLVLTLLNLFAGTETVSTTLRYGFL 315  
QY 227 -----HAATIAKKGGELLNMDYOKMEFTQCVISEALRCGNIVT-VHRRAT 273  
Db 316 LMKYVDIAKVEEIDRVIGKNNROKPFEDRKMPTEAVIHQIFADLIPKGLARV 375



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 20.97 Seconds  
(Without alignments) 1750.411 Million cell updates/sec

Title: US-09-995-917A-1  
Perfect score: 2027

Sequence: 1 MGMPFGETISFFKPHRSDS.....YVEFKRGMLEIETPKFLED 382

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: PIR71:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2018	99.6	382	2 T48613	hypothetical prote
2	821	40.5	513	2 T46143	steroid 22-alpha-h
3	655	32.3	472	1 S55379	cytochrome P450 CY
4	638	31.5	457	2 D85429	cytochrome P450 11
5	599.5	29.6	512	2 H96759	probable steroid 2
6	568.5	28.0	464	2 T07859	cytochrome P450 ho
7	504	24.9	457	2 T04444	cytochrome P450 -
8	487.5	24.1	485	2 A84859	probable cytochrom
9	484	23.9	482	2 T02739	probable cytochrom
10	460	22.7	455	2 T48973	cytochrome P450-11
11	432.5	21.3	490	2 H86185	hypothetical prote
12	427	21.1	489	2 B84733	probable cytochrom
13	416	20.5	487	1 C71417	cytochrome P450 d1
14	412.5	20.4	460	2 D96813	hypothetical prote
15	390.5	19.3	519	1 T02663	cytochrome P450 DM
16	370	18.3	255	2 T04602	cytochrome P450 ho
17	347.5	17.1	444	1 S75761	cytochrome P450 -
18	330.5	11.4	494	2 A33030	coumarin 7-monooxy
19	225.5	11.1	494	2 C34271	cytochrome P450 2A
20	223.5	11.0	494	2 A33531	testosterone 15alp
21	223	11.0	500	2 T04737	cytochrome P450 ho
22	219.5	10.8	494	2 B33531	cytochrome P450 2A
23	219	10.7	500	2 T52175	cytochrome P450 mo
24	217.5	10.6	494	2 S16068	testosterone 15alp
25	215.5	10.6	349	2 A86329	coumarin 7-monooxy
26	215.5	10.6	494	1 O4H0A6	FLAPL 4 protein -
27	215.5	10.6	494	1 O4H0A6	coumarin 7-hydroxy
28	206.5	10.2	494	2 A47494	cytochrome P450 2A
29	204.5	10.1	494	2 A33293	cytochrome P450 2A

## ALIGNMENTS

30	203	10.0	492	2	A34272	testosterone 7alp
31	199	9.8	492	2	S32491	testosterone 7alp
32	196.5	9.7	494	2	I38967	cytochrome P450 -
33	193	9.5	511	2	T00868	probable cytochrom
34	191.5	9.4	491	2	S31277	cytochrome P450 2B
35	191.5	9.4	494	2	I38965	cytochrome P450 -
36	190.5	9.4	492	2	S27160	cytochrome P450 2B
37	190	9.4	528	1	A31854	cytochrome P450 2B
38	189	9.3	501	2	T04735	cytochrome P450 ho
39	189	9.3	528	1	O4CK51	lanosterol 14alp
40	187.5	9.3	494	2	S11305	cytochrome P450 2B
41	185.5	9.2	491	1	O4RBP	cytochrome P450 2B
42	185.5	9.2	491	2	A27717	cytochrome P450 2B
43	185.5	9.2	501	1	A40938	cytochrome P450 2B
44	185	9.1	492	2	A31887	cytochrome P450 1b
45	184.5	9.1	522	2	T07118	testosterone 7alp
						probable cytochrom

ALIGNMENTS

RESULT	1	
T48613	hypothetical protein F18022.190 - Arabidopsis thaliana	
C:Species: Arabidopsis thaliana (mouse-ear cress)		
C:Date: 20-Apr-2000 #sequence-revision 20-Apr-2000 #text-change 20-Apr-2000		
C:Accession: T48613		
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.;		
submitted to the Protein Sequence Database, April 2000		
A:Reference number: Z24493		
A:Accession: T48613		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-382 <BEV>		
A:Cross-References: EMBL:DB		
A:Experimental source: cultured Columbia; BAC clone F18022		
C:Genetics:		
A:Map position: 5		
A:Introns: 31/2; 139/3; 204/1; 224/3; 253/3; 280/1; 315/3; 349/2		
A:Note: F18022.190		

Query Match	99.6%	Score 2018;	DB 2;	Length 382;	
Best Local Similarity	99.5%	Pred. No. 1e-145;			
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;					

QY	1	MGMPFGETISFFKPHRSDSITFLQOVRSGYKFKSNICGKAVSCDDELNNFIION	60
DB	1	MGMPFGETISFFKPHRSDSITFLQOVRSGYKFKSNICGKAVSCDDELNNFIION	60
QY	61	ECKLFTSDYPRKAMHDIIGKYSLLATNGEIRKLVVTSFINLTKSKDPLHCAENLSIS	120
DB	61	ECKLFTSDYPRKAMHDIIGKYSLLATNGEIRKLVVTSFINLTKSKDPLHCAENLSIS	120
QY	121	ILSKRNCREVEFEHKEVKNFTLSVAVNQLLSIKPEDPARLYLQDFLSYMGFISLPIPL	180
DB	121	ILSKRNCREVEFEHKEVKNFTLSVAVNQLLSIKPEDPARLYLQDFLSYMGFISLPIPL	180
QY	181	PCTGTNMIKVSNNIIONALIEDMNNAIREDPDLDSISNEDEHNAIRAKKDGELL	240
DB	181	PCTGTNMIKVSNNIIONALIEDMNNAIREDPDLDSISNEDEHNAIRAKKDGELL	240
QY	241	NMEDYQKMEFTQCVSEALRCGNIVKTVHRKATHDIKREYVTPKGMKVPFPTAVHIDP	300
DB	241	NMEDYQKMEFTQCVSEALRCGNIVKTVHRKATHDIKREYVTPKGMKVPFPTAVHIDP	300
QY	301	SIHENPFEFNPWRMTKTTAFGGGVAVCPGGELGKLQIAFLHLHLVLSYRWKIKSDMPIA	360
DB	301	SIHENPFEFNPWRMTKTTAFGGGVAVCPGGELGKLQIAFLHLHLVLSYRWKIKSDMPIA	360
QY	361	HPIYVEFKRGMLEIETPKFLED	382
DB	361	HPIYVEFKRGMLEIETPKFLED	382

RESULT 2  
T6143  
steroid 22-alpha-hydroxylase (DWP4) - Arabidopsis thaliana  
N:Alternate names: protein T3A5.40  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 16-Feb-2001  
C:Accession: T6143  
R:Blocker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetler, F.; Sala  
submitted to the Protein Sequence Database, December 1999  
A:Reference number: 223024  
A:Accession: T6143  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-513 <BL>  
A:Cross-references: EMBL:AL132979  
A:Experimental source: cultivar Columbia; BAC clone T3A5  
C:Genetics:  
A:Map position: 3  
A:Intons: 74/2; 182/3; 233/3; 358/3; 369/3; 396/1; 432/3  
A:Note: T3A5.40  
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
C:Keywords: heme; iron; metalloprotein  
F:308-484/Domain: cytochrome P450 homology <P45>  
F:462/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.5%; Score 821; DB 2; Length 513;  
Best Local Similarity 34.4%; Pred. No. 1.3e-54;  
Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;

QY 2 GMPFICTGFPPKRRSDSICGFLQQRVSRVGVKFKNSICGKAVVSCDDELNMFILQNE 61  
DB 45 GMPFLGFTGYLKYATYTGDFMQQVSKYGVKFKNSLNFPTVSADAGLRFILQNE 104  
QY 62 GKLTSDYPRKMHDLKYSLLATGEIHRKLVNIIISFINTLSKPDPLHCAENLSIS 121  
DB 105 GRLEECYPRSIGILGKMGMLVYGDHMDMSISLNFSHARLRLTILKDYERHTLFV 164  
QY 122 LKSKNCREVEPHKEVMEFLSVVNOGLSIKPEDPAKLYVLODFLSYMGFSLPLP 181  
DB 165 LDSWQQSISFSADEAKKFFENLAKHIMSDPEEETGLKREYVEMGVVASAPLNL 224  
QY 182 GTGYTNLIKVRNINIHQNIIEEDMNAIREDPFLDSIISNED----- 224  
DB 225 GTAVHKLQSRATILKFTIERKMERKIDIKEDDEEBEVKTEDAEKMSGDHYRKQRTD 284  
QY 225 -----EENAAI 230  
DB 285 DLGLMVLKHSNLTSEQLIDLLSLFAGHETSSVALAIFFLOACPKAVEELREHLEI 344  
QY 231 -RAKGGDEL-LNMDYQKMEFTOCVSEALRCGNIVKYTHRKATHDIKREYVTPGKW 288  
DB 345 ARAKKEGESELNWDYKKMDFTQVINETRLGNNVRFHRRKALKDYRKGYIDIPGKW 404  
QY 289 VFPIFTAVHDLPSLHNPFEFNPARKWTKT-----AFGGVAVPCG 329  
DB 405 VLPVTSANVHLDNSRKDQNPENLFPNRWQOONNGASSGSGSPSTWGNVYMPGGGPRLCAG 464  
QY 330 GELKQLQIAFLHVLVSYRKIKSDKEMPIAHVYERKGMLEI 374  
DB 465 SELAKLEMAVFIHVLKFNWELAEDDKPFAPFVDPNGIDIRV 509

RESULT 3  
S55379  
cytochrome P450 CYP90 - Arabidopsis thaliana  
N:Contains: oxidoreductase (EC 1.1.1.18)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: S55379  
R:Seckeres, M.; Nemeth, K.; Koncz, Z.; Nagy, F.; Koncz, C.

submitted to the EMBL Data Library, May 1995  
A:Reference number: S55379  
A:Accession: S55379  
A:Molecule type: mRNA  
A:Residues: 1-472 <SZE>  
A:Cross-references: EMBL:X87367; NID:g853718; PIDN:CAM60793.1; PID:g853719  
C:Genetics:  
A:Gene: CYP90  
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase  
F:275-440/Domain: cytochrome P450 homology <P45>  
F:418/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 32.3%; Score 655; DB 1; Length 472;  
Best Local Similarity 34.5%; Pred. No. 4.7e-42;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGPFIETISFPKRRSDSICGFLQQRVSRVGVKFKNSICGKAVVSCDDELNMFILQNE 60  
DB 36 LGPLIGEFPLIGAVKTEENPEFIDERVARYGVFMTHLFGEPTIFSADEPTNRFVLQ 95  
QY 61 EGKLTSDYPRKMHDLKYSLLATGEIHRKLVNIIISFINTLSKPDPLHCAENLSIS 120  
DB 96 EGKLECSYSPASICNLGKHSLLKMSLHKRHSILMSFANSITIKHMLMDIDRLVRF 155  
QY 121 LKSKNCREVEPHKEVMEFLSVVNOGLSIKPEDPAKLYVLODFLSYMGFSLPLP 180  
DB 156 NIDMSNS--RVLLMEPAKTIITFELYKOLMSFDPGWSSE-SLRKEYLVIEGFSLSPLP 212  
QY 181 PGTYTNLIKVRNINIHQNIIEEDMNAIREDPFLDSIISNED----- 224  
DB 213 FSTYTKAIQAR--RKVALALIVVVKRRREEEGEERKMDIALAADGFSDEEIVD 270  
QY 225 -----EENAAIRAKKGGDELNMDYQKME 249  
DB 271 FLVALIVAGETTTSTMTLAVKFLTEPLALQKREHKKIRAMSDSYSLMSDYKSNP 330  
QY 250 FTQCVISEALRCGNIVKYTHRKATHDIKREYVTPGKWVRFPIFTAVHDLPSLHNPFE 309  
DB 331 FTQCVNETHLVANIIIGVFRRAMTDEIKGKIPKGMKVFSSFRVHDLPSHFDARTF 390  
QY 310 NPKARKWT-----TAFGGVAVPCGDELKQLQIAFLHVLVSYRW 350  
DB 391 NPKRMQSNVVTGPNVFTFPGGPRLCPEYLARVALSVFLHRLVTGFSW 441

RESULT 4  
D85429  
cytochrome P450 like protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Apr-2001  
C:Accession: D85429  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; M0ID:20083488  
A:Accession: D85429  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-457 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7270586; PIDN:GAB80304.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g36380  
A:Map position: 4  
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
C:Keywords: heme; iron; metalloprotein  
F:396/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 31.5%; Score 638; DB 2; Length 457;  
Best Local Similarity 32.2%; Pred. No. 8.8e-41;  
Matches 141; Conservative 75; Mismatches 156; Indels 66; Gaps 7;



Db 331 RAVILETSLATLVNGVLRKTTODMEINGIIPKGRRIYYTRELINDEPLLYDPYSFNP 390  
 QY 312 MRWTKTT-----AFGGVAVCPGGELGKIQIAEFLHHLVLSYRKIKSDMPIAHPY 363  
 Db 391 WRWMDKSLHNGSLFVGGGTGCPGKELGVALISFLHFVFKYRWEEIGDGLMKFPR 450  
 QY 364 VEFKRGMLEI 374  
 Db 451 VEAPNGIRIRV 461

RESULT 7  
 T04444  
 cytochrome P450 - Arabidopsis thaliana  
 N:Alternate names: protein T18B16.200; protein T5K18.10  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 28-Jul-2000  
 C:Accession: T04444; T05806  
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H. M.  
 submitted to the Protein Sequence Database, April 1998  
 A:Reference number: Z15359  
 A:Accession: T04444  
 A:Molecule type: DNA  
 A:Residues: 1-457 <BEV>  
 A:Cross-references: EMBL:AL021687  
 A:Experimental source: cultivar Columbia; BAC clone T18B16  
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.T.; Voelt, M.; Robben, J.; Volckaert, G.; Ba  
 submitted to the Protein Sequence Database, April 1998  
 A:Reference number: Z15453  
 A:Accession: T05806  
 A:Molecule type: DNA  
 A:Residues: 131-457 <BEV>  
 A:Cross-references: EMBL:AL022580  
 A:Experimental source: cultivar Columbia; BAC clone T5K18  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 67/2: 173/3: 302/3: 358/1: 393/3  
 A>Note: T18B16.200; T5K18.10  
 C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
 C:Keywords: chromoprotein; electron transfer; heme; iron; monooxygenase; oxidoreductase  
 F:/272-433/Domain: cytochrome P450 homology <P45>

Query Match 24.9%; Score 504; DB 2; Length 457;  
 Best Local Similarity 29.3%; Pred. No. 1.3e-30;  
 Matches 120; Conservative 68; Mismatches 144; Indels 78; Gaps 11;  
 QY 1 MGMPFGETISFEPKPHRSDSITGLQORVSRGKGVKSNICGKAVVSCDQELNMFILQ 60  
 Db 41 MGMPYVEFQLY-----SDPNVFFQSKRGVSFETHVLGCPVMISPEAKFVLVT 96  
 QY 61 EKLFTSDYPRAMHDILGKYSLLATGEIHRKLNVI--SFT--NLTKSPDLFACENL 117  
 Db 97 KSHLFKTPFASKRMKGKQALFFHQDGHAKRLKVLRAFMPESIRNMVPI---ESI 152  
 QY 118 SISLTKMKKQREVEFPEKVMFTLSVNVNQLSI-----KPEDPARLYLVLODPLSYM 170  
 Db 153 AQSLSRWEGM--INTYQEMKTYFNV---ALLSIFGKDEVLYEDLKRCYIIE----- 203  
 QY 171 KGFSLPILPLPGTYNAIKVRN--RNHQNALIEDNMNNAIREDFLDSIISND----- 224  
 Db 204 KGYNSMNVNLPGLTFHKSMAKRELSQLIARILSERONGSHNDLLGSFWDKEELTDE 263  
 QY 225 -----EHAIRAKKGGLLNMDY 245  
 Db 264 QIADNIIGVIFPAHDTASVSMILKYLAENPNVLEAVTEQMAIRDKEGESLWGD 323  
 QY 246 QKMFETQCVSEALRCGNIYVTKRKAHDIKFEKVIYIPKGMKVFPIPTAVHILPSPSHEN 305  
 Db 324 KKMPLTRVIOETLRVASIISFTFREAVEDVEYGYILPKMKVPLPLFRNIHNSADIFSN 383  
 QY 306 PFEFPMRW-----TKTAFGGGVAVCPGGELGKIQIAEFLHHLVLSYR 349

Db 384 PGKFDPSREVPAPKPTNFMFPGNGTSCRGNELAKLEMSIHHHTTKR 433  
 RESULT 8  
 A84859  
 probable cytochrome P450 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 23-Mar-2001  
 C:Accession: A84859  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; M0ID:20083487  
 A:Accession: A84859  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-485 <STO>  
 A:Cross-references: GB:AE002093; NID:g4512670; PIDN:AAD21724.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g42850  
 A:Map position: 2  
 C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:/432/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 24.1%; Score 487.5; DB 2; Length 485;  
 Best Local Similarity 27.1%; Pred. No. 2.6e-29;  
 Matches 121; Conservative 79; Mismatches 163; Indels 83; Gaps 10;

QY 1 MGMPFGETISFEPKPHRSDSITGLQORVSRGKGVKSNICGKAVVSCDQELNMFILQ 59  
 Db 50 MGMPWGETMDYFKKQASNVFEDFVNPRLIKGNITKRTIMSPITVNGAENRLLIS 109  
 QY 60 NEGLFTSDYPRAMHDILGKYSLLATGEIHRKLNVI--ISFINLTKSRP-----DF 110  
 Db 110 NEPSLVVSMSPSSVQGLMGNICMAKGEKRRVLRGLVANSLSYIGLESILPKICDTYKF 169  
 QY 111 LHCAENLSISILSKWKNCREVEFPEKVMFTLSVNVNQLSIKPEDPARLYLVLODPLSYM 170  
 Db 170 HHETE-----WRKESEISYRSKAVLTFVVEFCLYIKVE---TGMLEVEFRLV 216  
 QY 171 KGFSLPILPLPGTYNAIKVR-----SNRNHONALIED 205  
 Db 217 EGVFALPVEPSCFARAKARLEIEFLVGYKREKREMEKGAEPNTTILSRVLEL 276  
 QY 206 MNNAIREDFLDSI-----SNED-----EHAIRAKKGG 237  
 Db 277 IKGVITEEVYDNNVLLVPAHDTTASVSMVTFKMLAQHPTCRDITLQEHANQIKANGEG 336  
 QY 238 ELLNMDYQKMEFTQCVSEALRCGNIYVTKRKAHDIKFEKVIYIPKGMKVFPIPTAVH 297  
 Db 337 EYLTVEDVKKMKYSWQVYRTRMLSPPIFSGFRAVADIDYGGTIPKGMKILMTWYGT 396  
 QY 298 LDPSLHENPEFNPMTWTKT-----TAFGGVAVCPGGELGKIQIAEFLHHLVLSYRWK 351  
 Db 397 YNPEIFQDPMSPDTRDKPLQATYILPFGGGRPLCAGHQLAKSILVFPHFVVTGDDWS 456  
 QY 352 IKSDEMPIAHPYVEF--RGMLEIEP 376  
 Db 457 LVYPDETISMDPLPFPGLGMPKISP 482

RESULT 9  
 T02739  
 probable cytochrome P450 At2g29090 [imported] - Arabidopsis thaliana  
 N:Alternate names: cytochrome P450 homolog T914.17  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001  
 C:Accession: T02739; D84692







QY 113 CAENLSTISILKSMKNCREVEFHKEKMTLSVYNQLL-SIKPEDPARLYVLQDFLSYMK 171  
 Db 147 DVLDTATYMEGARNGYLDVKTETSKILLGCLAKKWKGMEEBAKELALCKRY--FQS 204  
 QY 172 GFISLPRLPGTGT-----YTNA-----IKVSRNRNIHQ-- 199  
 Db 205 GMRFFFLNLPGTGVYKMKVLFVQYTEADISWQARKMKMLRKRYLTJKRASGEELGEFF 264  
 QY 200 NAIYEDM-----NNAIREDFLDLSISNE-----DEEH 227  
 Db 265 NIFGEGMEGEGEMSVENAV-EYIYFFELVANETTRILLATVKTISDHKKVQEOLOREH 323  
 QY 228 AAI---RAKGDDELNMEYQOMETQCVISALRCGNIVTKYHRAKTHIDIKFEYVIP 284  
 Db 324 EEIVRGRKAKEEGG--LTWEDYKSMHFTQWYNINSLRILISTAPYLVRLHEDFQVDYTIIP 381  
 QY 285 KGWKVPPIFTAVHLDPDSLHNPFEFPMRW-----TKT-TAGGVRVCPGSELGK 334  
 Db 382 AGW-TMGVYHIIHFNSEKEDPYAFNPMWREGDLCALIVSKTFIIPGAGRRLCVGAFAFK 440  
 QY 335 LQIAFFLHLVLSYRKIRKISDEMPHAPYVEFKRMILEI 374  
 Db 441 MQMAVFIHHL-FRYRWSMKSGTIIIRSFILMFGGCDVQI 479

RESULT 14  
 D96813  
 Hypothetical protein T30F21.17 (imported) - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: D96813  
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 Hansen, N.F.; Hughes, B.; Hultar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Matzfeld,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yi, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: D96813  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-460 <STO>  
 A:Cross-references: GB:AE005173; NID:94836883; PIDN:AD30586.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T30F21.17  
 A:Map position: 1

Query Match 20.4%; Score 412.5; DB 2; Length 460;  
 Best Local Similarity 27.4%; Pred. No. 1.2e-23;  
 Matches 121; Conservative 73; Mismatches 163; Indels 85; Gaps 13;  
 QY 1 MGPPIGETISFEKHSRSDSITGFLQOVRSGKFKFNICGGKAVVSCDDELNFILION 60  
 Db 40 MGFPIGETLDFKPCGVEGIFTFVKRMIRGPLEFTNIGSKIVVSTDDVYHQIRQ 99  
 QY 61 ECKLFTSDYKAMHDILKYSLLLATGEIHRKLVNIISFINLTKSRDPLHCAENLSIS 120  
 Db 100 EMTSPELGVDIFVAVFKDNLFLKEVFIHKYLOKITMOILGSEDLKATMGNMDKATRD 159  
 QY 121 ILSKKNCREVEFHKEVKMTLSVYNQLLS-IRKEDPARLY-VLQDF-LSYMKGFLSLP 177  
 Db 160 HIRSLASGSGFNRKVEVNLVAVYMTPLRISLTKPETOSKLIDNLAENLDMFKSFLRLS 219  
 QY 178 IPLPSTGTNAIKVRSNR-NIHONAIIEDMNAIREDFELDSISNEDE----- 225  
 Db 220 ---TKKATVTKALKSREELAIQVWKVDVLMRKREKQEDFLTLLEELKDGSGFDQGSAT 276  
 QY 226 -----EHAIRAKKGDEL-LNMEDYQ- 246

Db 277 NLIFLLAFLRECTSCTALAVKFSIKDPKVLAEIKREKRAIVDNKKKACGVSWEEVRH 336  
 QY 247 KMEFTQCVISEALRCGNIVTKYHRAKTHIDIKFK-EYVIK-----GMKFPPIFTAVHLDP 300  
 Db 337 NMFTNMAVSENEVLRLANTPLPLFRKAYQDVELIKKTYLEVYIHMGWL----- 384  
 QY 301 SIHENFEENPMRWTKTT--AFGGVRVCPGSELGKLQIAFFLHLVLSYRKIRKISDEMP 358  
 Db 385 -----WQKREIMWSKTFMAGYGVRLCVGAEFSSRLQMAIFLHLVAVYDFSMQDSEI 438  
 QY 359 IAHYVEFKRMILEI--EPTK 378  
 Db 439 IRSPFHQYTKDLINISQSPK 460

RESULT 15  
 T02263  
 Cytochrome P450 DNARF3 - maize  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Zea mays (maize)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: T02263  
 R:Winkler, R.G.; Helentjaris, T.  
 Plant Cell 7, 1307-1317, 1995  
 A>Title: The maize dwarf3 gene encodes a cytochrome P450-mediated early step in gibbe  
 A:Reference number: Z14648; MUID:96004534  
 A:Accession: T02263  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-519 <WIN>  
 A:Cross-references: EMBL:U32579; NID:9987266; PIDN:AA049067.1; PID:9987267  
 A:Experimental source: strain B73  
 A:Genetics:  
 A:Gene: dwarf3  
 A:Function:  
 A:Description: Involved in an early step in gibberellin biosynthesis  
 A:Pathway: gibberellin biosynthesis  
 A:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology  
 C:Keywords: oxidoreductase  
 F:325-488/domain: cytochrome P450 homology <P45>

Query Match 19.3%; Score 390.5; DB 1; Length 519;  
 Best Local Similarity 26.4%; Pred. No. 6.6e-22;  
 Matches 111; Conservative 69; Mismatches 160; Indels 81; Gaps 10;  
 QY 1 MGPPIGETISFEKHSRSDSITGFLQOVRSGK--VFKSNICGGKAVV-----SCDDEL 53  
 Db 77 MGWPLVGGMAFLRAFKSGKPDAFIASFVRGRGVRSFMSPTVLVTTAGCKQVL 136  
 QY 54 NMFILQNEGKLFSDYKAMHDILKYSLLLATGEIHRKLVNIISFINLTKSRDPLHLC 113  
 Db 137 -----HDDDAFVYGMKATYALVGRPSYVAMPYDEHRRIRKTLAAAINDEPDAITGLPF 190  
 QY 114 AENLSISILSM-NKREVEFHKEVKMTLSVYNQLLSTIKREDPARLYLQ--DFLSYMK 171  
 Db 191 IDRIVTSLRAMADHGSVEFLTELRLMTFKIIYQIFLG--GADQATRLALERSYTYLNY 248  
 QY 172 GFISLPRLPGTGTNAIKVRSNRNIHONAIIEDMNAIREDFELDSISNEDE-----DFLSISNE 224  
 Db 249 GMRMAINLPGFAIRGALRAR-RLVAVLOGVDERBARAKVSGGVGVMDBRLLEAOD 307  
 QY 225 -----EHAIRAKKG 235  
 Db 308 ERGRHLDDELDIVLVYLNAGHSSGHTIMATVFLQENPDMFAKAKADEAIMNSIPS 367  
 QY 236 DGEILNMEYQOMETQCVISEALRCGNIVTKYHRAKTHIDIKREYVIPGKVVPIFTA 295  
 Db 368 SQRCGLTLRDFRKMEYLSQVIDELRLVNIISFVSFRQATRDVFNGLVLPKGMVQVQWYRS 427  
 QY 296 VHLDPSSLHNPFEENPMRWTKTT-----AFGGVRVCPGSELGKLQIAFFLHLVLSY 348

Tue Jul 30 08:40:07 2002

us-09-995-917a-1.rpr

Page 8

Db 428 VHMPOVYPDPPTKFPDSRWEGHSPRAGTFLAIGLARGNDLAKLEISVFLHHFLIGY 487

OY 349 R 349

Db 488 K 488

Search completed: July 29, 2002, 13:58:04  
Job time: 31 sec



PS Claim 50; Fig 11; 113pp; English.

CC The present sequence represents a DMW4 polypeptide. The polypeptide is  
CC cytochrome P450 enzyme that mediates multiple steps in synthesis of  
CC brassinosteroids. Specifically, it mediates multiple  
CC 22alpha-hydroxylation steps in brassinosteroid biosynthesis. The DMW4  
CC polypeptide is used for altering the phenotype of a plant. DMW4  
CC plants display a dramatic reduction in the length of different organs,  
CC and this size reduction is attributable to a defect in cell elongation.  
CC The DMW4 polynucleotides and polypeptides can be used in diagnostic  
CC assays and to generate antibodies, which can be used to produce  
CC immunogenic compositions.

50	Sequence	513 AA
----	----------	--------

Query Match	40.6%	Score	822	DB	21	Length	513
Best Local Similarity	34.4%	Pred	No.	2	26-70		
Matches	160	Conservative	87	Mismatches	126	Indels	92
						Gaps	4

[illegible]

## RESULT 2

ID : AAG44572 standard; Protein; 444 AA

AC AAG44572;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55848.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.

OS *Arabidopsis thaliana*.

PN EP1033405-A2.

PD 06-SEP-2000.

PL XX

PF 25-FEB-2000; 2000EP-0301439.  
XX

PR	25-FEB-1999;	9905-0121825;
PR	05-MAR-1999;	9905-0123160;
PR	09-MAR-1999;	9905-0123548;
PR	23-MAR-1999;	9905-0125788;
PR	25-MAR-1999;	9905-0126264;
PR	01-APR-1999;	9905-0126785;
PR	01-APR-1999;	9905-0127462;
PR	06-APR-1999;	9905-0128224;
PR	08-APR-1999;	9905-0128714;
PR	16-APR-1999;	9905-0129845;
PR	19-APR-1999;	9905-0130007;
PR	21-APR-1999;	9905-0130549;
PR	23-APR-1999;	9905-0130510;
PR	28-APR-1999;	9905-0130891;
PR	30-APR-1999;	9905-0132048;
PR	04-MAY-1999;	9905-0132404;
PR	04-MAY-1999;	9905-0132484;
PR	05-MAY-1999;	9905-0132485;
PR	06-MAY-1999;	9905-0132486;
PR	07-MAY-1999;	9905-0132863;
PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134218;
PR	14-MAY-1999;	9905-0134219;
PR	14-MAY-1999;	9905-0134321;
PR	14-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134766;
PR	19-MAY-1999;	9905-0134941;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135353;
PR	25-MAY-1999;	9905-0135629;
PR	27-MAY-1999;	9905-0136392;
PR	28-MAY-1999;	9905-01363782;
PR	01-JUN-1999;	9905-01371222;
PR	03-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137503;
PR	07-JUN-1999;	9905-0137724;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138540;
PR	14-JUN-1999;	9905-0138847;
PR	14-JUN-1999;	9905-01393119;
PR	16-JUN-1999;	9905-01393452;
PR	16-JUN-1999;	9905-01393453;
PR	17-JUN-1999;	9905-0139492;
PR	18-JUN-1999;	9905-0138454;
PR	18-JUN-1999;	9905-0138455;
PR	18-JUN-1999;	9905-0139456;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139460;
PR	18-JUN-1999;	9905-0139461;
PR	18-JUN-1999;	9905-0139462;
PR	18-JUN-1999;	9905-0139463;
PR	18-JUN-1999;	9905-0139750;
PR	18-JUN-1999;	9905-01398173;
PR	21-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139899;
PR	23-JUN-1999;	9905-0140353;
PR	23-JUN-1999;	9905-0140354;
PR	24-JUN-1999;	9905-0140695;
PR	28-JUN-1999;	9905-0140823;
PR	29-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	02-JUL-1999;	9905-0142154;
PR	06-JUL-1999;	9905-0142055;
PR	08-JUL-1999;	9905-0142290;

PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 18-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 655; DB 21; Length 444;  
Best Local Similarity 34.5%; Pred. No. 2,2e-54;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGHPFGETISFFPKRRSSIGTFLOQRVSRGVKFSNICGKAVVSCDQELMFIION 60  
DB 8 l9pllgctfqlgkxktempepfideravarysvfemthlfgeptlfssdpctnrfvqpn 67  
QY 61 EGRLETSDBPKAMHDITKGVSLTATGETHRKRLKXNITFIINLTSKRPFLCAENLTS 120  
DB 68 egklfecsybastcnllghsalllmkxslhkrmhslmfsfanssllkdhlmldrlvrf 127  
QY 121 IKSWMKREVEFHKKEKMTLSVMVNDLSIKPEDPARLVYLODFLSYMKGFISLPIPL 180  
DB 128 nldswes--rvllmeekklfelfcvkqlmsfdpgewse-slrkeyllviegffslppl 184  
QY 181 PGTGTNAIKVSNRNRIHONAI-----EDMNAIREDPDLISITNED----- 224  
DB 185 fctlyrkaigdar--rkvaallvvyvkrtreeeegarkkdmllaadddfsdeetv 242  
QY 225 -----EENAIIRAKKGDDGELLMWEDYOKME 249  
DB 243 flvallvagyettstclmclavkflecplalaglkveehetkrtamsdsyslswsdykmp 302  
QY 250 FTQCVIISALRCGNIVKTVHRRATHDIKREYVIRKGMKVPPIPTAVHLDPSLHNPPEF 309  
DB 303 ftcgvvneltrvanllgyfrramcdveikykrlpkgykvwfssfravnlrdphfkdarlf 362  
QY 310 NPMRMTKT-----TARGGVRCVPGGSLGKLQIAFLFHNHLVSYRW 350  
DB 363 nptwgsnsvltgpnsvftptg9gprlcpgyelarvalvalsvflhrlvtgfs 413

RESULT 3  
AAG45023  
ID AAG45023 standard; Protein: 444 AA.

XX AAG45023;  
AC 18-OCT-2000 (first entry)  
DT  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56470.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138004.  
PR 10-JUN-1999; 99US-0138340.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.



```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 32.3%; Score 655; DB 21; Length 444;  
 Best Local Similarity 34.5%; Pred. No. 2.2e-54;  
 Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

OY 1 MGPWFIGETISFFKPHRSDSICGTFLQDVRVRYGVKFSNICGKAVVSCDDELNMFIION 60
DB 8 Iglpligtctfqlgkytkenpepfidervarygsvfmhlgpeplifadpctrnrfvqgn 67
OY 61 ECKLTSTVPRAMHDIKYSILLATGFIHRKLNKVVITSFNLTGSKPDPLCAENLSTIS 120
DB 68 egkltccyspsicnllgkhhllmkgsllkrmshlmsfnsllkdhlmldldtrlyrf 127
OY 121 ILKSWKNCREVEFHKEVMEFTLSVWVNOULSIKPEDPARLVLDPEFLSYMGFTSLPPL 180
DB 128 nldwsvs---rrllmeeakkifeltvkdqmsfdpgeuse-sirkeyllivagfslppl 184
OY 181 PGTGTNAIKVRSNRNHIQNAII-----EDMNNAIREDPLDLSITSNED----- 224
DB 185 fstlyrkaiqr--fkvealtlvvmkrreeegaeerkkdmlaallaaaddgfsdeetld 242

```

```

OY 225 -----EBHAIRAKKGDGELNWDYQKWE 249
DB 243 flvaillvagyeltstmltavkflteclajlaqlkeekhrakmsdsyalswsgykamp 302
OY 250 FTQCVISEALRCGNIVKTVHRAKTHDIKREYVIRKGMKVPFPIFAVHLDPSLHNPEEF 309
DB 303 ftgcgvneelirvanllgvyfirramcdvclkykldpkykvtssfvavhldphfkdarlf 362
OY 310 NPMRWTKT-----TAFGGVAVCPGCELGKIQIAFFLHLVLSYRW 350
DB 363 npwrwgsnsvltgpnvtfpfggprlcpgyelarvalsvflhrlvtgfsw 413

RESULT 4
AAW27153
ID AAW27153 standard; Protein; 472 AA.
XX
AC AAW27153;
XX
DT 14-APR-1998 (first entry)
DE Arabidopsis thaliana cytochrome P450-type hydroxylase.
XX
KW Cytochrome P450-type hydroxylase; identification; brassinosteroid;
KW brassinosteroid inhibitor; modified plant; recombinant production;
KW testosterone.
OS Arabidopsis thaliana.
XX
PN W09J35986-A1.
XX
PD 02-OCT-1997.
XX
PE 27-MAR-1997; 97WO-EP01586.
XX
PR 27-MAR-1996; 96US-0622166.
XX
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI Altmann T, Koncz C, Mathur J, Szekeres MA;
XX
XX WPI: 1997-489649/45.
XX DR N-PSDB; AAT85306, AAT85307.
XX
PT New isolated plant cytochrome P450-type hydroxylase gene - used to
PT identify substances acting as brassino-steroid(s) or brassinosteroid
PT inhibitors for the production of modified plants
XX
PS Claim 1; Pages 44-46; 77pp; English.
XX
XX The present sequence is Arabidopsis thaliana cytochrome
XX P450-type hydroxylase. The hydroxylase can be used to identify
XX brassinosteroids or brassinosteroid inhibitors, useful to produce
XX plants with modified physiological and/or phenotypic
XX characteristics. The modified plants may show, e.g. stimulation of
XX growth, increased seed elongation, increased wood production,
XX accelerated seed germination at low temperatures, an increase in
XX dry weight, repressed anthocyanin production during growth in light
XX and/or inhibited de-etiolation which is induced, e.g. by cytokinin,
XX in the dark or an increase in stress tolerance. The hydroxylase or
XX its coding sequence can also be used for the recombinant production
XX of compounds, e.g. testosterone.
XX
SO Sequence 472 AA:

```

Query Match 32.3%; Score 655; DB 18; Length 472;  
 Best Local Similarity 34.5%; Pred. No. 2.4e-54;  
 Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

OY 1 MGPWFIGETISFFKPHRSDSICGTFLQDVRVRYGVKFSNICGKAVVSCDDELNMFIION 60
DB 185 fstlyrkaiqr--fkvealtlvvmkrreeegaeerkkdmlaallaaaddgfsdeetld 242

```

Db	36	lgdrlpdlgetfqligaykntenpepfidervarygsvfmthlfgpecltfdsadpetntrfvldn	95		
Oy	61	EGKRLFTSDYRKAMMDILGKYSLLLATGELIRKLNVIISFINTLTKSPDFLHCQENLSTIS	120		
Db	96	egkrlfteesypasicnllgkhsllllmkgsllhkmhslltmefanssllkdhlmldidrlvrf	155		
Oy	121	ILKSWKACREVEYFEKFKMTFLSMVNOQLSLIRPEDPARLYVQDPLSLMKGFISLPIPL	180		
Db	156	nldswss--rvllmeekaklftelvtvkqmsfdpgewse-slrkeyllvlegifslpdl	212		
Oy	181	PGCTGYTAIKVRSNRNTHQNAII-----EDMNNALIREDFLDSITSNED-----	224		
Db	213	fstyrkaiagr--rkvaeatlvvmkrreeeegaerkkdmlaalaaaddgfsdeeiwd	270		
Oy	225	-----EENHAIRAKKGDGELLNEDYQKME	249		
Db	271	flvalllvagyettclmtlawnkflteclplaaqlkeehexikremksdyslswdykssmp	330		
Oy	250	FTQCYSIALSCGNIVKTVHKKATHDIFKEEYVLPKGMKVFPIFTAVHLDPSLHNPFFER	309		
Db	331	ftqgcvmctlrvanligvfrfamtvdclkykpkqwkxfsfrefvahlidpnhfkdarft	390		
Oy	310	NPMRWTKT-----TAFGGSVRVCPGGELCKLOIAFLHLHLVLSYRW	350		
Db	391	npwrqgsnsvtqpsnvftfpfggprlrcpgyelarvalswflnrlvtcgisw	441		
RESULT 5					
ID	AAg44571	AAg44571 standard; Protein: 472 AA.			
AC	AAg44571;				
DT	18-OCT-2000	(first entry)			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55847.				
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.				
OS	Arabidopsis thaliana.				
PN	EP1033405-A2.				
XX	06-SEP-2000.				
XX	25-FEB-2000; 2000EP-0301439.				
PR	25-FEB-1999;	99US-0121825.			
PR	09-MAR-1999;	99US-0123180.			
PR	09-MAR-1999;	99US-0122548.			
PR	23-MAR-1999;	99US-0125788.			
PR	25-MAR-1999;	99US-0126264.			
PR	29-MAR-1999;	99US-0126785.			
PR	01-APR-1999;	99US-0127462.			
PR	06-APR-1999;	99US-0128234.			
PR	08-APR-1999;	99US-0128714.			
PR	16-APR-1999;	99US-0128845.			
PR	19-APR-1999;	99US-0130077.			
PR	21-APR-1999;	99US-0130449.			
PR	23-APR-1999;	99US-0130510.			
PR	23-APR-1999;	99US-0130891.			
PR	28-APR-1999;	99US-0131449.			
PR	30-APR-1999;	99US-0132048.			
PR	30-APR-1999;	99US-0132407.			
PR	04-MAY-1999;	99US-0132484.			
PR	05-MAY-1999;	99US-0132485.			
PR	06-MAY-1999;	99US-0132486.			
PR	07-MAY-1999;	99US-0132487.			
PR	11-MAY-1999;	99US-0132863.			
PR	14-MAY-1999;	99US-0134256.			
PR	14-MAY-1999;	99US-0134218.			

PR	14-MAY-1999;	9905-0134219;
PR	14-MAY-1999;	9905-0134221;
PR	14-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134768;
PR	19-MAY-1999;	9905-0134941;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135353;
PR	24-MAY-1999;	9905-0135629;
PR	25-MAY-1999;	9905-0136021;
PR	27-MAY-1999;	9905-0136382;
PR	01-JUN-1999;	9905-0136782;
PR	08-JUN-1999;	9905-0137222;
PR	03-JUN-1999;	9905-0137502;
PR	07-JUN-1999;	9905-0137724;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138840;
PR	14-JUN-1999;	9905-0139119;
PR	16-JUN-1999;	9905-0139452;
PR	16-JUN-1999;	9905-0139453;
PR	17-JUN-1999;	9905-0139492;
PR	18-JUN-1999;	9905-0139454;
PR	18-JUN-1999;	9905-0139455;
PR	18-JUN-1999;	9905-0139456;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139460;
PR	18-JUN-1999;	9905-0139461;
PR	18-JUN-1999;	9905-0139462;
PR	18-JUN-1999;	9905-0139463;
PR	18-JUN-1999;	9905-0139750;
PR	18-JUN-1999;	9905-0139763;
PR	21-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139889;
PR	22-JUN-1999;	9905-0140035;
PR	23-JUN-1999;	9905-0140059;
PR	24-JUN-1999;	9905-0140659;
PR	28-JUN-1999;	9905-0140823;
PR	29-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142390;
PR	06-JUL-1999;	9905-0142905;
PR	08-JUL-1999;	9905-0144803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144085;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144332;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144884;
PR	21-JUL-1999;	9905-0144881;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;

```

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0144386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160991.
PR 22-OCT-1999; 99US-0160999.
PR 22-OCT-1999; 99US-0160999.

```

```

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%, Score 655; DB 21; Length 472;
Best Local Similarity 34.5%, Pred. No. 2,4e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGPPIGETISFEKPHRSDSIGFLQQRVSRGKVKFNSICGKAVVSCDDELNNFIQON 60
   :| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 1gplrlgetfglqgyktenpepfidervarysvfmhlfgeptlfadpetnrfvign 95

QY 61 EGGLEFTSDYPRKAMHDILCKYSILLATGEIHRKKNVITISFTNLKSKDFLHCANLSIS 120
   ||||| | : : : : ||||| | : : : : | : : : : | : : : : | : : : : |
Db 96 egklfecsyaspasilqkhsillmkgselhkrmhsilmsfmslsikdhlmdldirvrf 155

QY 121 ILKSWKNCREVEFHKVEKFTLSVKNQLSIRKPEDPARLVLODFLSYMKFSLPIPL 180
   | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 156 nldsws--rvllmeekklfelftvkqlmsfdpgwse-sireyallvgeftslppl 212

QY 181 PGTGYTNAIKVRSNBNHQNIAIT-----EDMNNVAREDFLDITSINED----- 224
   | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 213 fstlyrkalgat--rkveaalvvvmkrreeegaeerkkdmllaaladddgfadeelvd 270

QY 225 -----EEHAIRAKKGCGLLNMEDYKME 249
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 271 flvalvagyettstlmclavkflteplalaqlkeeneklramksdyslsesdysmp 330

QY 250 FTQCVISEALRCGNIVKTVHRKATHDIKFEKYEYIPKGVKVPFTAAHLDPSLHENPEEF 309
   ||||| : | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 331 ftgcgvnecltvanllgfvframtdevelkgyklpkqwkvsfsfrvahlpnhfkdrctf 390

QY 310 NPMRWTKT-----TARGGVRCVPGGELCKLQIDAFPLHLVLSYRW 350
   || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 391 npwtwgsnsvltgpsnvftplg9gprlpcgyelarvalsvflhrlvctgfsaw 441

RESULT 6
AAG45022
ID AAG45022 standard; Protein; 472 AA.
XX
AC AAG45022;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56469.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

```

PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130049.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138099.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139899.  
PR 22-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0144390.  
PR 08-JUL-1999; 99US-014803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 27-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157173.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.

PR	13-OCT-1999:	9905-0155294.
PR	13-OCT-1999:	9905-0155295.
PR	14-OCT-1999:	9905-0155329.
PR	14-OCT-1999:	9905-0155330.
PR	14-OCT-1999:	9905-0155331.
PR	14-OCT-1999:	9905-0155637.
PR	14-OCT-1999:	9905-0155638.
PR	18-OCT-1999:	9905-0155684.
PR	21-OCT-1999:	9905-0160741.
PR	21-OCT-1999:	9905-0160747.
PR	21-OCT-1999:	9905-0160767.
PR	21-OCT-1999:	9905-0160768.
PR	21-OCT-1999:	9905-0160770.
PR	21-OCT-1999:	9905-0160814.
PR	21-OCT-1999:	9905-0160815.
PR	22-OCT-1999:	9905-0160980.
PR	22-OCT-1999:	9905-0160981.
PR	22-OCT-1999:	9905-0160989.
PR	25-OCT-1999:	9905-0161404.
PR	25-OCT-1999:	9905-0161405.
PR	25-OCT-1999:	9905-0161406.
PR	26-OCT-1999:	9905-0161359.
PR	26-OCT-1999:	9905-0161360.
PR	26-OCT-1999:	9905-0161361.
PR	28-OCT-1999:	9905-0161920.
PR	28-OCT-1999:	9905-0161992.
PR	28-OCT-1999:	9905-0161993.
PR	29-OCT-1999:	9905-0162142.

Query Match	32.38;	Score 655;	DB 21;	Length 472;
-------------	--------	------------	--------	-------------

```

Best Local Similarity 34.58; Pred. NO. 2.4e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

[illegible]

XX	RESULT	7
XX	AA045021	ID
XX	AA045021	standard; Protein; 491 AA.
XX	AA045021;	
XX	18-OCT-2000	(first entry)
XX	Arabidopsis thaliana protein fragment	SEQ ID NO: 56468.
XX	Protein identification:	signal transduction pathway; metabolic pathway;
XX	hybridisation assay;	genetic mapping; gene expression control; promoter

KM		termination sequence.
XX		
OS		<i>Arabidopsis thaliana</i> .
XX		EPI033405=-f2.
PN		
XX		
PD		06-SEP-2000.
XX		
PF		25-FEB-2000; 2000EP-0301439
XX		
PR	25-FEB-1999;	99US-0128125
PR	05-MAR-1999;	99US-0123180
PR	23-MAR-1999;	99US-0125748
PR	23-MAR-1999;	99US-0125358
PR	25-MAR-1999;	99US-0126264
PR	29-MAR-1999;	99US-0126785
PR	01-APR-1999;	99US-0127462
PR	06-APR-1999;	99US-0128234
PR	08-APR-1999;	99US-0128714
PR	16-APR-1999;	99US-0129845
PR	19-APR-1999;	99US-0130077
PR	21-APR-1999;	99US-0130449
PR	23-APR-1999;	99US-0130510
PR	23-APR-1999;	99US-0130891
PR	28-APR-1999;	99US-0131448
PR	30-APR-1999;	99US-0132048
PR	30-APR-1999;	99US-0132407
PR	04-MAY-1999;	99US-0132484
PR	05-MAY-1999;	99US-0132882
PR	06-MAY-1999;	99US-0132487
PR	06-MAY-1999;	99US-0132486
PR	07-MAY-1999;	99US-0132863
PR	11-MAY-1999;	99US-0134256
PR	14-MAY-1999;	99US-0134218
PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134421
PR	14-MAY-1999;	99US-0134370
PR	18-MAY-1999;	99US-0134376
PR	19-MAY-1999;	99US-0134541
PR	20-MAY-1999;	99US-0135124
PR	21-MAY-1999;	99US-0135553
PR	24-MAY-1999;	99US-0135629
PR	25-MAY-1999;	99US-0136021
PR	27-MAY-1999;	99US-0136592
PR	28-MAY-1999;	99US-0136782
PR	01-JUN-1999;	99US-0137222
PR	03-JUN-1999;	99US-0137528
PR	04-JUN-1999;	99US-0137620
PR	07-JUN-1999;	99US-0137724
PR	08-JUN-1999;	99US-0138094
PR	10-JUN-1999;	99US-0138540
PR	10-JUN-1999;	99US-0138847
PR	14-JUN-1999;	99US-0139119
PR	16-JUN-1999;	99US-0139452
PR	16-JUN-1999;	99US-0139493
PR	17-JUN-1999;	99US-0139493
PR	18-JUN-1999;	99US-0139454
PR	18-JUN-1999;	99US-0139455
PR	18-JUN-1999;	99US-0139463
PR	18-JUN-1999;	99US-0139750
PR	18-JUN-1999;	99US-0139763
PR	21-JUN-1999;	99US-0139817
PR	22-JUN-1999;	99US-0139899
PR	23-JUN-1999;	99US-0140353
PR	24-JUN-1999;	99US-0140354
PR	24-JUN-1999;	99US-0140695

KM		termination sequence.
XX		
OS		<i>Arabidopsis thaliana</i> .
XX		EPI033405=-f2.
PN		
XX		
PD		06-SEP-2000.
XX		
PF		25-FEB-2000; 2000EP-0301439
XX		
PR	25-FEB-1999;	99US-0128125
PR	05-MAR-1999;	99US-0123180
PR	23-MAR-1999;	99US-0125748
PR	23-MAR-1999;	99US-0125358
PR	25-MAR-1999;	99US-0126264
PR	29-MAR-1999;	99US-0126785
PR	01-APR-1999;	99US-0127462
PR	06-APR-1999;	99US-0128234
PR	08-APR-1999;	99US-0128714
PR	16-APR-1999;	99US-0129845
PR	19-APR-1999;	99US-0130077
PR	21-APR-1999;	99US-0130449
PR	23-APR-1999;	99US-0130510
PR	23-APR-1999;	99US-0130891
PR	28-APR-1999;	99US-0131448
PR	30-APR-1999;	99US-0132048
PR	30-APR-1999;	99US-0132407
PR	04-MAY-1999;	99US-0132484
PR	05-MAY-1999;	99US-0132882
PR	06-MAY-1999;	99US-0132487
PR	06-MAY-1999;	99US-0132486
PR	07-MAY-1999;	99US-0132863
PR	11-MAY-1999;	99US-0134256
PR	14-MAY-1999;	99US-0134218
PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134421
PR	14-MAY-1999;	99US-0134370
PR	18-MAY-1999;	99US-0134376
PR	19-MAY-1999;	99US-0135441
PR	20-MAY-1999;	99US-0135424
PR	21-MAY-1999;	99US-0135553
PR	24-MAY-1999;	99US-0135629
PR	25-MAY-1999;	99US-0136021
PR	27-MAY-1999;	99US-0136592
PR	28-MAY-1999;	99US-0136782
PR	01-JUN-1999;	99US-0137222
PR	03-JUN-1999;	99US-0137528
PR	04-JUN-1999;	99US-0137620
PR	07-JUN-1999;	99US-0137724
PR	08-JUN-1999;	99US-0138094
PR	10-JUN-1999;	99US-0138540
PR	10-JUN-1999;	99US-0138847
PR	14-JUN-1999;	99US-0139119
PR	16-JUN-1999;	99US-0139452
PR	16-JUN-1999;	99US-0139493
PR	17-JUN-1999;	99US-0139493
PR	18-JUN-1999;	99US-0139454
PR	18-JUN-1999;	99US-0139455
PR	18-JUN-1999;	99US-0139463
PR	18-JUN-1999;	99US-0139750
PR	18-JUN-1999;	99US-0139763
PR	21-JUN-1999;	99US-0139817
PR	22-JUN-1999;	99US-0139899
PR	23-JUN-1999;	99US-0140353
PR	24-JUN-1999;	99US-0140354
PR	24-JUN-1999;	99US-0140695

PR	28-JUN-1999;	9905-0140983;
PR	29-JUN-1999;	9905-0140981;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	02-JUL-1999;	9905-0142350;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144884;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	22-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	28-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147360;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	10-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149173;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149900;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	25-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151503;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;

PR	10-SEP-1999;	9905-01530770;
PR	13-SEP-1999;	9905-01533758;
PR	15-SEP-1999;	9905-015404018;
PR	16-SEP-1999;	9905-015404039;
PR	20-SEP-1999;	9905-01544779;
PR	22-SEP-1999;	9905-01551139;
PR	23-SEP-1999;	9905-01551486;
PR	24-SEP-1999;	9905-01556559;
PR	28-SEP-1999;	9905-01564658;
PR	29-SEP-1999;	9905-01565966;
PR	04-OCT-1999;	9905-01571711;
PR	05-OCT-1999;	9905-01571753;
PR	06-OCT-1999;	9905-01578665;
PR	07-OCT-1999;	9905-01586029;
PR	08-OCT-1999;	9905-01586233;
PR	12-OCT-1999;	9905-01583669;
PR	13-OCT-1999;	9905-01585293;
PR	13-OCT-1999;	9905-01592294;
PR	14-OCT-1999;	9905-01593295;
PR	14-OCT-1999;	9905-01593320;
PR	14-OCT-1999;	9905-01595331;
PR	14-OCT-1999;	9905-01595637;
PR	14-OCT-1999;	9905-01595688;
PR	18-OCT-1999;	9905-01595841;
PR	21-OCT-1999;	9905-01607671;
PR	21-OCT-1999;	9905-01607667;
PR	21-OCT-1999;	9905-01607668;
PR	21-OCT-1999;	9905-01608174;
PR	21-OCT-1999;	9905-01608815;
PR	21-OCT-1999;	9905-01608815;
PR	22-OCT-1999;	9905-01609811;
PR	22-OCT-1999;	9905-01609814;
PR	22-OCT-1999;	9905-01609893;
PR	25-OCT-1999;	9905-01614004;
PR	25-OCT-1999;	9905-01614005;
PR	25-OCT-1999;	9905-01614065;
PR	26-OCT-1999;	9905-01613559;
PR	26-OCT-1999;	9905-01613660;
PR	26-OCT-1999;	9905-01613661;
PR	28-OCT-1999;	9905-01616920;
PR	28-OCT-1999;	9905-01616920;
PR	29-OCT-1999;	9905-01621442;
PR	29-OCT-1999;	9905-01621493;

Query Match	32.3%	Score 655;	DB 21;	Length 491;
-------------	-------	------------	--------	-------------

Best Local Similarity 34.58; Pref. No. 2.6e-54;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```

Qy 1 MOWPIGEFISPFKPHRSDS:IGTFLOODSVKSYGVFNSJGCAVWSCQOELMNTLON 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 jdlrlagetfdq:lgayktempepfidervaygsvmthlfgpeltlsadpentrflqn 114

Qy 61 ECKLFTSDYPKAMHDLGKYSLLLATGEHRLKKNVILSFNLTKSPPDLHCAENLMS 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 egklicecspasicollqkshlilmkslhrkrmstlmsfansiikdhmlaidetlvrf 174
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 IKSMSKNCREYEFHKEVMEFLTSVWNLDSIRKPEDPARLVLODFLSYMKGCTSPILP 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 nldsnss--rylmeeaakktifeltvtvqmsfdpgense-slrkegllvlegffs:plprl 231

Qy 181 PGTGYTNAIKVRSNBNHONAI-----EDMNNAIRDEDFLDSITSND 224
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 fatyfrkaigar-fkvaaalvvumkrreeeegaeraktkdmalaallaaadgfsdeeiwd 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 225 -----EHHAIIRAKKGDEGLIMMEDYOKME 249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 flvallvagyettstlmclavkflfelpblaaqlkeehetlrmksdssylswdsykmp 349

Qy 250 PTCVASEALRCGNLIVTKHRRKATHDIKEKEYVTIPKCKWFPRTAYVHDLPSLHENPFEF 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 ffcgvavnehlrvanlignvframtdveibkyk:lpkwwkssfs:travhaldpnmfkaartf 409

```



PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 655; DB 21; Length 492;  
Best Local Similarity 34.5%; Pred. No. 2,6e-54;  
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGWFIQETISFRPHRSISIGFPLQORVSRYGKVRKSNICGKAVYSCQELMFILO 60  
Db 56 lglpligetqlgkyktenpepidervarysvfmthlfigeptlfsadpetnrtfvlgn 115  
QY 61 EGRFTSDYPRAMHDIIGKXSLILATGEIHRKILKNVTSFLNLTJKSPDFLHCACENLSIS 120  
Db 116 egklfecsyastcnllqkshllmkyslhmrmstmsfamsiikdhlmldtrivrf 175

QY 121 ILSWKNCREVEFHKEVKMFTLSVMYNOLLSTIKEDPARLVYLDELFSYMKGPISLPIPL 180  
Db 176 nldswas--fvlimeekakltfelvkgqmsfdpgewse-slrkeylllydegffsfpjl 232  
QY 181 PGGTYNARKVRNSRNRIHQNAII-----EDMNNAIREDFDLDSIISND----- 224  
Db 233 fsttyrtaidgr--rkvaaeltvvmkrrreeegearckdmiaallaadgfsdeelv 290  
QY 225 -----EENHAIRAKKGDGELLMEDYQKME 249  
Db 291 flvalliwaqyetslmltlavkfltlepllaqlkeehkeklramksdyslswsykmp 350  
QY 250 FTQCVISEALRCGNIVKTVRKATHDIKEKEYVYIPKGWVFPIFTAVHLDPSLHENPFEF 309  
Db 351 ftgcvmeltrvanliigvtrramtdvelkyklpkykwkfsfravhdpnhfkdarlf 410  
QY 310 NPMRWRT-----TAFGGGVRCVCPGEGELQIAEFFLHLVLSYRW 350  
Db 411 nprwqnsvtlqpsnvftpfqg9pripcyelarvalsvflhrlytgisw 461

RESULT 9  
ID AAG11836 standard; Protein: 461 AA.  
XX AAG11836;  
AC AAG11836;  
XX 17-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10708.  
DE  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0133484.  
PR 05-MAY-1999; 99US-0133485.  
PR 06-MAY-1999; 99US-0133486.  
PR 06-MAY-1999; 99US-0133487.  
PR 07-MAY-1999; 99US-0133863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.



PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0135621.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145219.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148568.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153785.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155113.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.



	PR	19-JUL-1999;	99US-0144333.
	PR	19-JUL-1999;	99US-0144334.
	PR	19-JUL-1999;	99US-0144335.
	PR	20-JUL-1999;	99US-0144332.
	PR	20-JUL-1999;	99US-0144632.
	PR	20-JUL-1999;	99US-0144884.
	PR	21-JUL-1999;	99US-0144814.
	PR	21-JUL-1999;	99US-0145087.
	PR	21-JUL-1999;	99US-0145088.
	PR	22-JUL-1999;	99US-0145085.
	PR	22-JUL-1999;	99US-0145087.
	PR	22-JUL-1999;	99US-0145089.
	PR	23-JUL-1999;	99US-0145129.
	PR	23-JUL-1999;	99US-0145145.
	PR	23-JUL-1999;	99US-0145218.
	PR	26-JUL-1999;	99US-0145224.
	PR	27-JUL-1999;	99US-0145276.
	PR	27-JUL-1999;	99US-0145913.
	PR	27-JUL-1999;	99US-0145919.
	PR	28-JUL-1999;	99US-0145919.
	PR	02-AUG-1999;	99US-0145921.
	PR	02-AUG-1999;	99US-0145931.
	PR	02-AUG-1999;	99US-0146386.
	PR	02-AUG-1999;	99US-0146388.
	PR	03-AUG-1999;	99US-0146389.
	PR	03-AUG-1999;	99US-0147038.
	PR	04-AUG-1999;	99US-0147204.
	PR	04-AUG-1999;	99US-0147302.
	PR	05-AUG-1999;	99US-0147132.
	PR	05-AUG-1999;	99US-0147260.
	PR	06-AUG-1999;	99US-0147303.
	PR	06-AUG-1999;	99US-0147416.
	PR	09-AUG-1999;	99US-0147493.
	PR	09-AUG-1999;	99US-0147495.
	PR	10-AUG-1999;	99US-0148171.
	PR	11-AUG-1999;	99US-0148319.
	PR	12-AUG-1999;	99US-0148341.
	PR	13-AUG-1999;	99US-0148565.
	PR	13-AUG-1999;	99US-0148564.
	PR	16-AUG-1999;	99US-0149368.
	PR	17-AUG-1999;	99US-0149175.
	PR	18-AUG-1999;	99US-0149426.
	PR	20-AUG-1999;	99US-0149722.
	PR	20-AUG-1999;	99US-0149723.
	PR	23-AUG-1999;	99US-0149929.
	PR	23-AUG-1999;	99US-0149929.
	PR	23-AUG-1999;	99US-0149930.
	PR	25-AUG-1999;	99US-0150566.
	PR	26-AUG-1999;	99US-0150884.
	PR	27-AUG-1999;	99US-0151085.
	PR	27-AUG-1999;	99US-0151085.
	PR	27-AUG-1999;	99US-0151086.
	PR	30-AUG-1999;	99US-0151303.
	PR	31-AUG-1999;	99US-0151438.
	PR	01-SEP-1999;	99US-0151930.
	PR	07-SEP-1999;	99US-0152363.
	PR	10-SEP-1999;	99US-0153070.
	PR	13-SEP-1999;	99US-0153758.
	PR	13-SEP-1999;	99US-0154018.
	PR	16-SEP-1999;	99US-0154039.
	PR	20-SEP-1999;	99US-0154779.
	PR	22-SEP-1999;	99US-0155139.
	PR	23-SEP-1999;	99US-0155486.
	PR	24-SEP-1999;	99US-0155659.
	PR	28-SEP-1999;	99US-0156559.
	PR	29-SEP-1999;	99US-0156458.
	PR	04-OCT-1999;	99US-0156596.
	PR	05-OCT-1999;	99US-0157117.
	PR	06-OCT-1999;	99US-0157753.
	PR	07-OCT-1999;	99US-0157865.
	PR	07-OCT-1999;	99US-0158029.
	PR	08-OCT-1999;	99US-0158332.
	PR	12-OCT-1999;	99US-0158359.
	PR	13-OCT-1999;	99US-0159293.
	PR	13-OCT-1999;	99US-0159294.
	PR	13-OCT-1999;	99US-0159294.
	PR	13-OCT-1999;	99US-0159295.
	PR	14-OCT-1999;	99US-0159320.
	PR	14-OCT-1999;	99US-0159330.
	PR	14-OCT-1999;	99US-0159331.
	PR	14-OCT-1999;	99US-0159637.
	PR	14-OCT-1999;	99US-0159638.
	PR	18-OCT-1999;	99US-0159584.
	PR	21-OCT-1999;	99US-0160741.
	PR	21-OCT-1999;	99US-0160767.
	PR	21-OCT-1999;	99US-0160770.
	PR	21-OCT-1999;	99US-0160814.
	PR	21-OCT-1999;	99US-0160815.
	PR	22-OCT-1999;	99US-0160980.
	PR	22-OCT-1999;	99US-0160981.
	PR	22-OCT-1999;	99US-0160989.</

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX EP103405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134770.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 22-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140635.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151303.

```

PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 28.2%; Score 572; DB 21; Length 465;

Best Local Similarity 30.7%; Pred. No. 2,4e-46; Indels 84; Gaps 9;

```

Matches 136; Conservative 71; Mismatches 152; Indels 84; Gaps 9;

QY 1 MGMPFICETISFFKPHRSDSICGTFLOQRVSRGKVFKSNICGKAVVSCDOELNMFIION 60
DB 39 mgwprfgtctefllkgpdr-----lmkngrlllygseffkshllgcpitvsmdehlnryllm 93
QY 61 ECKLFTSDYPRKAMHDILGKYSLLATGEIHRKLNVIISFINTLTKSRKDFLHCAENLSIS 120
DB 94 eskjlvagypqsmidllgltcnlaevhgsphrlmrgsllslslspcmkdhllpkidlmrn 153
QY 121 ILKSKKNCREVEFHKVEMFLLSVNQLLSI-----KPE-DPAKLYLVQDFLSYMGKFI 174
DB 154 ylgcwddletvdigektkhm--afllslqlaetlkkpeveeyr---teffkflvgtl 206
QY 175 SLPLPCTGYTNAIKVNSRNHONAIIEDMNNAIRE-----DFLDSIISNED---- 224
DB 207 svpdlipgtnyrgsvqanmldrllltelmg-----tkesgettdmlylmkkehdyrl 261
QY 225 -----EENHAIRAKKGDGELIN 241
DB 262 ltkeltdqvcllysgyelvtctsmmaljyldhpkaleelrrehlaierkrpdeplc 321
QY 242 MEDYQKMEFTQCVISSEALRCGNIVKTVHRKATHDIKFEKVIYIPKGMVPIFTAVHLDPS 301

```

```

DB 322 lddksmkftravlfetrsrlatlngvrltkthdlnghyllpkywrylvylreindts 381
QY 302 LHENPEFENPMRKTKT-----AFGGVAVCPGCGELGKLOIAFPFLHLVLSYRMKIK 353
DB 382 lyedpmifnpwrmekslesksyfillfggyvlrcpckelglsavsfihyftkyreen 441
QY 354 SDEPIAHPIYVEFKRGMLEIEP 376
DB 442 gedklmfvprvsapkyhkcsp 464

```

#### RESULT 12

AAU02839 standard; Protein: 486 AA.

AAU02839;

07-SEP-2001 (first entry)

Taxus cuspidata oxygenase enzyme #19.

Oxygenase; Japanese yew; Taxol; taxoid; Taxol biosynthetic pathway;

transgenic organism.

Taxus cuspidata.

WO200134780-A2.

17-MAY-2001.

13-NOV-2000; 2000WO-US31254.

12-NOV-1999; 9905-0165250.

(UNIW ) UNIV WASHINGTON STATE RES FOUND.

Croteau RB, Schoendorf A, Jennewein S;

WPI: 2001-355489/37.

DR N-PSDB; AAS05196.

Novel nucleic acid and amino acid sequences, isolated from the Taxus

genus, useful for the synthetic production of Taxol and related

taxoids, intermediates within the Taxol biosynthetic pathway, and

other taxoid derivatives -

Claim 1; Page 139-141; 141pp; English.

The sequence represents a Taxus cuspidata oxygenase, encoded by a DNA of the invention. The sequences isolated from the Taxus genus, and the respective oxygenases are useful for the synthetic production of Taxol and related taxoids, as well as intermediates within the Taxol biosynthetic pathway, as well as intermediates within the Taxol can be used to make transgenic organisms that either produce the oxygenases for subsequent in vitro use, or produce the oxygenases in vivo so as to alter the level of Taxol and taxoid production within the transgenic organism. The oxygenase nucleic acids and amino acids are useful for isolating the polynucleotide and polypeptide sequences corresponding to full-length oxygenases.

Sequence 486 AA;

Query Match 27.5%; Score 557; DB 22; Length 486;

Best Local Similarity 30.9%; Pred. No. 7.1e-45;

Matches 126; Conservative 64; Mismatches 158; Indels 60; Gaps 8;

```

QY 2 GMPFICETISFFKPHRSDSICGTFLOQRVSRGKVFKSNICG-KAVVSCDOELNMFIION 60
DB 52 glpfigetlftldakspgrrkffidelnrlgprlcsilgrrtravsvdpefnkyvlqn 111
QY 61 ECKLFTSDYPRKAMHDILGKYSLLATGEIHRKLNVIISFINTLTKSRKDFLHCAENLSIS 120

```



```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```

```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 546; DB 21; Length 461;
Best Local Similarity 29.7%; Pred. No. 7,66-44;
Matches 130; Conservative 73; Mismatches 101; Indels 74; Gaps 9;

OY 1 MCMPIGETISEFPHRSDSISGTFLQDRVSRGKVKFNSICGRAVNSCDLNMFILOK 60
DB 35 mgywlfgeteteflkggn-----fmnqrltygsftekhllygcpellsmdeavnyllkn 89
OY 61 ECKLFTSDPKAMHDILGKYSILLATGEIHRKLVVITSPINLTKSKDPFLHCENLS 120
DB 90 eekgllvpyqgsmldllglcmaavhgashtlmrgslslsstmrmchllpkvdhmr 149
OY 121 LKSMKNCREVEFHKEVKMFTLSVWVNOULS--IKPEDPARLYVLQ--DFLSYMKGFIS 175
DB 150 yldqmelevidgdktkmaflslltqaghlrpk-----fveefktaffklvvgls 203
OY 176 LPPLPGTGYTNAIKVRNSRNHONAIIDMNA-----IREDDF 215
DB 204 vpldlpgtnyrcgldgarnldtflrelmqerrdsgetfdmlygmkegnrypltdael 263
OY 216 LDSIIS-----NEDEHNAIRKKDGDGLNMEDYOK 247
DB 264 rdqavtlllysgyelvstsmma[ky]hdpkalgelraehlafrckrqdelpgldevks 323
OY 248 MEFTQCVISEALRCGNIYVTHRKATHDIKFEKEYIIPKGMKVPPIPTAVHLDPSLHNP 307
DB 324 mkftravlyetrlatlngvrlrkttrdelngyllpkwrylvvrtrengdanlyedpl 383
OY 308 EENPMKWTKT-----AFGGVAVCPGCGELGKIQTAFELHVLVLSRW-KISDENP 358
DB 384 lfnpwrmkkslæspscfvfggtrtlcpgekylveissflhyftrlyweelsgdel- 442
OY 359 IAHPYVEFKRGMLETER 376
DB 443 mvfprvfapkgfhlrlisp 460

RESULT 14
AAG46489
ID AAG46489 standard; Protein; 462 AA.
XX
XX AAG46489;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58493.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX
XX 06-SEP-2000.
PE
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

```

PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135114.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0143270.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151530.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.





```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153785.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 546; DB 21; Length 465;
Best Local Similarity 29.7%; Pred. No. 7,7e-44;
Matches 130; Conservative 73; Mismatches 161; Indels 74; Gaps 9;

OY 1 MGVWPIGFTISFHKHRSDSIGTFLOQRVSRGKVKESNIGCAVSCDOLNFIION 60
    ||||| ||||| | : : : ||||| : : : | : | : | : | : | : |
Db 39 mgwplfgetelkxgpr-----lmnqrlltygsfksnllgcprllismdevmrlylkn 93

OY 61 ECKLFTSDYPRKAMHDILDKYSLLATGEIHRKLKNVSIISFTNLTKSKDPFLHCAENLIS 120
    | : | ||||| | : : : | : : | : : | : : | : : | : : |
Db 94 eskglvpyppysgmldlbtcmavhsgshrlmrgslslslsttmrthljpkydhfms 153

OY 121 IUKSKNCREVEFHKVKMFTLSVWNOULLS-IRKEDPARLYUO---DPLSYKGFIS 175
    | : | : : : : : : : : : : : : : : : : : : : : : : :
Db 154 yldqwelevldlqdktkhmafslsltqjagnlrkp-----fveefktafkflvgtls 207

OY 176 LPIPLPGTYTAIKVRSRNIHQNAIIEDNMNA-----IREEDF 215
    : : | ||||| | : : : | : : : : : : : : : : : : : : :
Db 208 vpldipgtnyrcigqarnldrllelmgerrdsgetltdmlylmkvegmrlypltddei 267

OY 216 LDSIIS-----NEDEHAAIRAKKGDGELNMEDYOK 247
    | : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 rdqavtlllysgyetvstcsmmaikylhdpkpalgelirehlafrerkrqdeplgjedvks 327
```



**THIS PAGE BLANK (uspro)**



PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151308.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157717.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	23-OCT-1999;	9905-0161404.
PR	23-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

Query Match	9.4%	Score 107.6	DB 21	Length 1667
Best Local Similarity	56.1%	Pred. No. 5.9e-19		
Matches 240	Conservative 0	Mismatches 164	Indels 24	Gaps 1
Qy	675	agaacatcagcgcattatgagcccaagaagaagggtggtgggagacttttggaattgggaagata	734	
Db	1002	agaacacttgcctacaaggagagaagaacgacctgacggaacgcctcaactctcgagatat	1061	
Qy	735	tcaagaagatgatcatcactcaatgtgtgattcttctgagagcactacgaatgtgtgaataatcgt	794	
Db	1062	tbaatcgaatgaattcaccacccgcgcgtgtgctcttggagatcagaattgtgcaacgattgt	1121	
Qy	795	caagactgtacatagaagaagaagcctaccatcatatattaatccaagaatatgtattccaaa	854	
Db	1122	tbaatgtgtcctcttaggaataactacacgactcagaactcagaacgcggtattattatcccaaa	1181	
Qy	855	gggtgtgaaagggtgttccatctctccagcagatcactcttgatccctctctcatgaaaa	914	
Db	1182	aggtctggagaatttaagcttataccagaagaagattaaactatgatactcctcttatgaga	1241	
Qy	915	tcccttggaaatbtaatcccatagaaatgagccaacga-----	952	
Db	1242	tccaatgatccttaaccccaatgagatgagatgagaaagagcttgaataccaagagctattt	1301	
Qy	953	--cggagcttggaggggaggtgaagggtatgctctgtgtgtgaactctggcagagctccaaat	1010	
Db	1302	ctbactccttggaggtgaggtcttaggtcttgccctggaaaagaaactcggaaatcccggaagt	1361	
Qy	1011	tgtctctctccctcacatctgtccctctccctaaagttggaataaagtcaagatgaat	1070	
Db	1362	ctaaagctctccctcactactctgttacaanaatagatggaagagaatgtgagaagacaa	1421	
Qy	1071	gccaaatcg 1078		
Db	1422	attaatcg 1429		
RESULT 2				
AAT85306				
ID AAT85306 standard; cDNA to mRNA. 1608 BP.				
XX AAT85306;				
XX				
DT 14-APR-1998 (first entry)				
XX Arabidopsis thaliana cytochrome P450-type hydroxylase cDNA.				
DE				
XX Cytochrome P450-type hydroxylase; identification: brassinosteroid.				
XX				

KW	brassinosteroid inhibitor; modified plant; recombinant production;
KW	testosterone; ds.
XX	
OS	Arabidopsis thaliana.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	48.1467
XX	/tag= a
XX	/product= cytochrome_p450-type_hydroxylase
PN	W09735986-A1.
XX	
PD	02-OCT-1997.
XX	
PE	27-MAR-1997; 97WO-EP01586.
XX	
PR	27-MAR-1996; 96US-0622166.
XX	
PA	(PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX	
PI	Altmann T, Koncz C, Mathur J, Szekeeres MA;
XX	WPI. 1997-469649/45.
DR	P-PSDB; AAW27153.

XX New isolated plant cytochrome P450-type hydroxylase gene - used to  
 PT identify substances acting as brassino-steroid(s) or brassinosteroid  
 PT inhibitors for the production of modified plants  
 XX  
 XX  
 PS Claim 1; Pages 44-46; 77pp: English.  
 XX  
 XX The present sequence encodes *Arabidopsis thaliana* cytochrome  
 CC P450-type hydroxylase. The hydroxylase can be used to identify  
 CC brassinosteroids or brassinosteroid inhibitors, useful to produce  
 CC plants with modified physiological and/or phenotypic  
 CC characteristics. The modified plants may show, e.g. stimulation of  
 CC growth, increased cell elongation, increased wood production,  
 CC accelerated seed germination at low temperatures, an increase in  
 CC dry weight, repressed anthocyanin production during growth in light  
 CC and/or inhibited detritolation which is induced, e.g. by cytokinin,  
 CC in the dark or an increase in stress tolerance. The hydroxylase or  
 CC its coding sequence can also be used for the recombinant production  
 CC of compounds, e.g. teasterone.  
 CC  
 SQ Sequence 1608 BP; 382 A; 374 C; 374 G; 478 T; 0 other;  
 XX

[illegible]

### RESULT 3

AAC48157  
ID AAC48157 standard; DNA: 1646 BP.  
XX AAC48157;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56467.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
FE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 29-MAR-1999; 99US-0126264.  
PR 01-APR-1999; 99US-0126785.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137228.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.



```
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155639.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 8.5%; Score 97.8; DB 21; Length 1646;  
Best Local Similarity 58.9%; Pred. No. 2,6e-16;  
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```
QY 671 atgaagaacatgacgcattagagcaagaagaagatggagggaacttgaattggaag 730
Db 972 aggaagaagcatgaaagaattagggcaatgaagaagctatcgtatagcttgaatggagtg 1031
QY 731 atatacagaagatggaattcaactcaaatgctggaatttcgaggcactacgattggttaata 790
Db 1032 attacaaagcttaacgcatcaccacaaatgctggttaatgagcgtacgagtcgcttaaca 1091
QY 791 tcgtcaagaactgtacatagaagaagcactacatgatataatcaagaatatgtgattc 850
Db 1092 tcactcgcggtgtcttccaaagctgcaatgacgagatgctgagatcaagaattataaatc 1151
QY 851 caaagggtggaagtgcttccaatcttcacagcagatcatctgtatccctctcttcatg 910
```

```
Db 1152 caaaagggtgaaagatattctcatcgttttagagcggttcatcttagaccnaaccacttca 1211
QY 911 aaatccttgaatttaaccatggatggagaccaaaagagg 955
Db 1212 aagatgtcgcacttcaacccttggagatgagcagagcaactcgg 1256

RESULT 4
AAC47986
ID AAC47986 standard; DNA; 1649 BP.
XX
AC AAC47986.
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55845.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131448.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
```

PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0136244.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145144.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145916.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157153.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 14-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 8.5%; Score 97.8; DB 21; Length 1649;  
Best Local Similarity 58.9%; Pred. No. 2.6e-16;  
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
Qy 671 atgaagaacatgcagccattagagccaaagaaggatggggaacttttggaattgggaag 730  
Db 974 aggaagaagcatgaagaagattgggcaatgaagaagtgtatagttcttgatgtgaatggagtg 1033

Oy 721 attacagaagatggaattcaattcaatgaatgtgattctctgaggcactacgaatgtgtaata 790  
Db 1034 attacaagatcgaatgcattcacacaaatgtgtgtaaatgagcgtacgagtggtctaa 1093  
Oy 791 tgcgaagaacgtgatactgaaaagctactcctgaattcaattcaagaataatgtgattc 850  
Db 1094 tcaatcgagcggtgtcttctcgaagcgtgaatgacgagatgttgagatcacaaggtataaaattc 1153  
Oy 851 caaaagggtcgaagggtgtcttccaaatctcacagcagatcaacttgaatccctctcatg 910  
Db 1154 caaaagggtcgaagggtatcttccaaatctcacagcagatcaacttgaatcccaacttca 1213  
Oy 911 aaaaaccttgaattcaattccatgagatggaacaaacgacgg 955  
Db 1214 aagatgctgcacattcaaccttggagatgagcagagaactcgg 1258

RESULT 5  
AAC48714  
ID AAC48714 standard; DNA: 1398 BP.  
XX  
AC AAC48714;  
XX  
DT 18-Oct-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58492.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130049.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132885.  
PR 06-MAY-1999; 99US-0132886.  
PR 06-MAY-1999; 99US-0132887.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141852.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147204.

```

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      8.4%; Score 96.8; DB 21; Length 1398;
Best Local Similarity 53.0%; Pred.No. 4,5e-16;
Matches 249; Conservative 0; Mismatches 197; Indels 24; Gaps 1.

QY 676 gaacatgcacgcattagacgaagaaggagatggygaactttgtaattggaagaattat 735
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 916 gagcatttggcattccagggaagaagaacgacagacgaaccactcgctcttgagacggt 975

QY 736 cagaagatggaattcaactaatgtgtgatttctgagacactcagatggttaattcgctc 795
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 976 aagtcattgaagttcactcgcagctgattttagagacataagattggaacgacgctgt 1035

QY 796 aagacgttcataagaaagcactcatgatattaaattcaagaatattgattccaaag 855
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1036 aatggggtccaaagaaactactcgttgcattggaatacaacggttatttaaccacaaa 1095

QY 856 ggggtggaaggtgtttccaaacttcacagcagttacatcttgatccctcttcataaatt 915
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1096 ggatggagaattattgtatatacagcagggaaatttaactgaatcttatttgaagac 1155

QY 916 cctttgaatttaaccatgaattgaccacaaagca----- 952
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1156 ccaattgactttaaccatgagatgagatgaagaagccttgagatcacaaaactatgc 1215

QY 953 -cgagcgtttgagagagagtaaggatgtccctggtgtgacttggcgaagctccaaatt 1011
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1216 ttgtgttttggaggttggacaagccttgcctgtaaggaactaggtatgcagatc 1275

QY 1012 gcttcttcctcaatcattgtctctcctcctatagttggaataaagtaagttgaatg 1071
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1276 tcgagcttctcattcattctgttgcagagatacagatgaggaagtaagagggatgaa 1335

QY 1072 ccaatgcgcacccttaacggtgagtttaagagagagcagcttggagat 1121
   || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1336 ttaatggtgttccgagagtttttgcaccaaaggcttccatttgat 1385

RESULT 6
AAS05196
ID AAS05196 standard; DNA; 1461 BP.
XX
AC AAS05196;
XX
DT 07-SEP-2001 (first entry)
XX
DE Taxus cuspidata oxygenase enzyme DNA #19.
XX
KW Oxygenase; Japanese yew; Taxol; taxoid; ds; Taxol biosynthetic pathway;
   transgenic organism.
XX
OS Taxus cuspidata.
XX
PN W0200134780-A2.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US31254.
XX
PR 12-NOV-1999; 99US-0165250.
XX
PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI Croteau RB, Schoendorf A, Jennewein S;
XX
DR WPI: 2001-355489/37.
XX
P-PSDB: AAU02839.
XX
PT Novel nucleic acid and amino acid sequences, isolated from the Taxus
   genus, useful for the synthetic production of Taxol and related
   taxoids, intermediates within the Taxol biosynthetic pathway, and
   other taxoid derivatives -
XX

```





RESULT 9  
AAC39029  
ID AAC39029 standard; DNA; 1685 BP.  
XX  
AC AAC39029;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23101.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PM EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 27-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 6.2%; Score 71.2; DB 21; Length 1685;  
Best Local Similarity 49.4%; Pred. No. 3.8e-09;  
Matches 230; Conservative 0; Mismatches 218; Indels 18; Gaps 1;

QY 636 aaaaagtttcgtgattcgaatattcgaatgaagatgaacatgcagccattgagc 695  
DB 1002 acatgtgacgaacacttcctaagaagctgttaaggtcgaagaaagctatatatgaaga 1061  
QY 696 caaagaaggagatgaggaaactttggaattggaagaattatcagaagatggaattcactca 755  
DB 1062 aaaaagctagagagaagaacacttaacatgagagacaacgaggaatattgcactgacaca 1121  
QY 756 atgtgtgatttcgagagcaactacgattgtgtaatatatcgtccaagactgttcataagaagc 815

DB 1122 taagtgtatagttgaagcttgagatgagcaagcatcatatcccttcacatcgaagaagc 1181  
QY 816 tactcatgtatataattcaagaatatgttatccaaagggtgtgaagttttccaat 875  
DB 1182 agtgtgtgtgtgataataaggaatattgatacccaaggaagtggaagtatgcact 1241  
QY 876 cttaacagcagttacatccttgatccctctctcatgaaatcctttgaaatccaat 935  
DB 1242 gtttcggaatatcatcaacaatccgaatatattttcaaacctgaggttttcgaccacac 1301  
QY 936 gagatg-----gaccaaacgcagcgcttggaggaagtaaggt 977  
DB 1302 tagattcgaagtaaatccgaagccgaatatcatcattgccttttggaagtggagttcatgc 1361  
QY 978 atgtcctgtgtgtaactctggcaagctccaattgtcttcctcatcattcgttcct 1037  
DB 1362 ttgtcccggaagcaactctgcgaagttacaattcttatattcttcaccattagttc 1421  
QY 1038 ctccctaaagtggaataaagtcagatgaatgccaatcgcgac 1083  
DB 1422 caattccgagtgggaagtgaaggaggaagaagaataacagttac 1467  
RESULT 10  
AAC37912  
ID AAC37912 standard; DNA; 940' BP.  
XX  
AC AAC37912;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19104.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW Protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP103405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.



PR 19-MAY-1999;	99US-0139441.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 14-JUN-1999;	99US-0138847.
PR 16-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 17-JUN-1999;	99US-0139453.
PR 18-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0143542.
PR 13-JUL-1999;	99US-0143624.
PR 14-JUL-1999;	99US-0144005.
PR 15-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 26-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147004.
PR 04-AUG-1999;	99US-0147208.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 09-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 10-AUG-1999;	99US-0147935.
PR 11-AUG-1999;	99US-0148171.
PR 12-AUG-1999;	99US-0148319.
PR 13-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 16-AUG-1999;	99US-0148684.
PR 17-AUG-1999;	99US-0149368.
PR 18-AUG-1999;	99US-0149175.
PR 20-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 23-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 25-AUG-1999;	99US-0149930.
PR 26-AUG-1999;	99US-0150566.
PR 27-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 30-AUG-1999;	99US-0151080.
PR 31-AUG-1999;	99US-0151303.
PR 01-SEP-1999;	99US-0151438.
PR 07-SEP-1999;	99US-0151930.
PR 10-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0156458.
PR 28-SEP-1999;	99US-0156596.
PR 29-SEP-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157733.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.



```

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145212.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145291.
PR 28-JUL-1999; 99US-0145291.
PR 28-JUL-1999; 99US-0145291.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```

```

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160781.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 5.9%; Score 67.8; DB 21; Length 1673;
Best Local Similarity 54.0%; Pred. No. 3.1e-08;
Matches 181; Conservative 0; Mismatches 127; Indels 27; Gaps 1;

```

```

QY 727 gaagattatcagaagatgaattcaactaatgtgtatcttgagggcactcagatgtgt 786
    ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 gaatatagacacacatgacttccacacatggtttagcaatgaggtgtcgtcgtgca 1151

QY 787 aatcgtcgaagactgtcatagaaagctactatgatatcaatcaagaagatgtg 846
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1152 aacacgaccccttgtgttctcgaagcggtcgaagatgttgagatcaagaagataca 1211

QY 847 attccaaggggtggaaggttttccaacttccaagcagatcctgtatccctctt 906
    |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 1212 attcggctgtgtgtatgtgtgtcggttcacacttcagcggtcatcttgcctgcaatc 1271

QY 907 catgaatactcttgaattcaatccatcgatgag----- 942
    |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 1272 tatgagaacccattgagtttaactcatgagatgagggagaaatgattgagga 1331

QY 943 ---accaaacgagcggttggagagatgaaggtatgtcctgtgtgtaacttggc 999
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1332 tctaaacgtcatcggttgcgttgatagatgaggttgaacttgtgtgagcagagtttcg 1391

QY 1000 aagctccaatgtcttcttccatcatcatctgt 1034
    || |||| | | |||| | | |||| | | |||| | | |||| | | |||| | | ||||
Db 1392 cggtccaatgtcaatcttccatcatcatctgt 1426

```

```

RESULT 12
AAC42516
ID AAC42516 standard; DNA; 1648 BP.
XX AC AAC42516;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35853.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.

```

PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148177.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.

```

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      5.2%; Score 59.6; DB 21; Length 1648;
Best Local Similarity 49.1%; Pred. No. 5e-06;
Matches 207; Conservative 0; Mismatches 194; Indels 21; Gaps 1;

```

```

QY 654 gataattcgaatgaagatgaagaaacatctagagcgaagaaaggatgagga 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1138 gattctgcnaaagcgttaagaaagagagatcgtlaaanaagagagacccggaca 1197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 714 actcttgaaattggaaagattatcagaagaatgaaatcactcaatgtgtatctgagagc 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1198 gaagtgaagcgttaaagagacacgtgaatcgtgtatcttctcaggttatgtatgagac 1257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 774 actacgaatgtgtaatactcgtcaagaactgtacataagaaagctactcgtatataatt 833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1258 ccttcgagtaattacatctctctcgtacggccttcagggagaaagaaatgtatgtccaat 1317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 834 caaagaataatgtatctcaaaagggtggaaggtgtttccaaatctcacagcagtaactc 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1318 ggaatgcataataatcnaaaggctggaaggtcttgacttggtttaagaaagctcattt 1377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 tgatccctctctcctcgaagaaatcctttgaatttaattccatgaaatgaaacaaagac 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1378 ggaacctgaatctcaccggtacaaagaatctgattccttcaaatgaggaggaataac 1437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 934 ggcg-----tttggagaggaagtaagggtatgtccctgtgtgtga 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1438 accaaagcagcagacatctcctctctgttggagagcacaactatgccgggaaagaa 1497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 993 acttggcagatcccaatgtcttctctcctcaatcattgtctctcctataaggtgaa 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1498 tcttgcagaagctcgaatctcatttcttcaatcattctcctcctcaataacggtgga 1557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1053 aa 1054
    ||
DB 1558 aa 1559

```

```

RESULT 13
AAS05171
ID AAS05171 standard; DNA; 1503 BP.

```

```

XX XX
AC AAS05171;
XX 07-SEP-2001 (first entry)
DT XX
XX Taxus cuspidata oxygenase enzyme DNA #6.
DE XX
XX Taxus cuspidata oxygenase enzyme DNA #6.
XX KM
XX Oxygenase; Japanese yew; Taxoid; taxoid; ds; Taxol biosynthetic pathway;
XX Transgenic organism.
XX Taxus cuspidata.
OS
XX WO200134780-A2.
XX PN
XX 17-MAY-2001.
XX PD
XX 13-NOV-2000; 2000WO-US31254.
XX PF
XX 12-NOV-1999; 99US-0165250.
XX PR
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX PA
XX Croteau RB, Schoendorf A, Jennewein S;
XX PI
XX WPI: 2001-355489/37.
XX DR
XX P-PSDB; AAU02826.
XX XX
XX Novel nucleic acid and amino acid sequences, isolated from the Taxus
PT genus, useful for the synthetic production of Taxol and related
PT taxoids, intermediates within the Taxol biosynthetic pathway, and
PT other taxoid derivatives -
XX
PS Claim 4; Page 99-100; 141pp; English.
XX
XX The sequence represents a DNA which encodes a Taxus cuspidata oxygenase.
CC The sequences isolated from the Taxus genus, and the respective
CC oxygenases are useful for the synthetic production of Taxol and related
CC taxoids, as well as intermediates within the Taxol biosynthetic pathway,
CC and other taxoid derivatives. The sequences also can be used to make
CC transgenic organisms that either produce the oxygenases for subsequent
CC in vitro use, or produce the oxygenases in vivo so as to alter the level of
CC Taxol and taxoid production within the transgenic organism. The oxygenase
CC nucleic acids and amino acids are useful for isolating the polynucleotide
CC and polypeptide sequences corresponding to full-length oxygenases.
XX
SQ Sequence 1503 BP; 406 A; 331 C; 324 G; 441 T; 1 other;

```

```

Query Match      5.1%; Score 59; DB 22; Length 1503;
Best Local Similarity 45.0%; Pred. No. 7.1e-06;
Matches 264; Conservative 0; Mismatches 320; Indels 3; Gaps 1;

```

```

QY 2 tggagatggccttcaattggagaactattctcttcaaaacctcalagatcagactca 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 tagctctcccttcaattggaggagacacatacattcttgaggcacttcgtcagaagacac 235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 tcggtatctcttgcaacaagctgtttccacggtatcggaagaaagtgttaagtcaaatat 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 ctacgaagtttttgatgagaaggtggaagaatctcgtgtgtatctcaagactcgtatag 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 122 ttggttggaagaagcagtagtctcatgtgaccaaagaaacctcaacatgtctataactcaaacg 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 ttgggcatcccaacagttgtactctcgcggtcggtggaagaaacgcttctctcctcaacg 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 aagggaagtgttataatcgtatatacnaaagcagatgacatctcgcgaataatc 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 aggaacaactggtgcaagcgltcatltgccaactcttcggaagaaactaatgtggaataatc 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 cccctcattagcaccgcggaagaattcacaggaacaaataaagtatataatagctca 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 416 ccattctgtccaaaagggaggagacatcgcatatataagtgctgtcaactgccgcgtt 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 tcaatctacaagaatcgaaaactgtacttctcaacttgcagagagaacctctatctga 361

```



```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155386.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 4.6%; Score 52.4; DB 21; Length 660;
Best Local Similarity 48.1%; Pred. No. 0 00032;
Matches 149; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

```

```

QY 1 atggatgagccttcattggaagaactattcttctcaaacctatagatcagctcc 60
DB 224 atggttaccgatactcgagagacatgcgactcttcttgagcccatgattatccg 283
QY 61 atcggatcattcttcgaacaacggtgttcacggtatggaagaagtggttcaagtc 120
DB 284 atccacaccttgtaagaagaagatgattaaagtgagggccattgttctgcaaacatc 343
QY 121 tgtgtgaaagaacagtagctcatgttgacccaagaactcaacatgttcaactcaaac 180
DB 344 ttcgataccaacacccgtgttctgacagaagctgatacatcttgaagtttccgcaa 403
QY 181 gaagggaagtgtttacatcgattatccaaaagcgtacatgacattctcgcaaatat 240
DB 404 gagaacaagcttcttgactagctataccagagccattctcaagcatttggaaaaga 463
QY 241 tccctctatagccacccggagagaattccaggaactaaagaattttatgcttc 300
DB 464 acgtgttccctcaacaacatggaacatccacaagcagtcacaacatcagttcaacttc 523
QY 301 atcaatctca 310
DB 524 ttgctctga 533

```

```

RESULT 15
AAS05174
ID AAS05174 standard; DNA; 1494 BP.
XX
AC AAS05174;
XX
DT 07-SEP-2001 (first entry)
XX
DE Taxus cuspidata oxygenase enzyme DNA #9.
XX
KW Oxygenase; Japanese yew; Taxol; taxoid; ds; Taxol biosynthetic pathway;
KW transgenic organism.
XX
OS Taxus cuspidata.
XX
PN WO200134780-A2.
XX
PD 17-MAY-2001.
XX
PE 13-NOV-2000; 2000WO-US31254.
XX
PR 12-NOV-1999; 99US-0165250.
XX
PA (UNIM ) UNIV WASHINGTON STATE RES FOUND.
XX
PI Croteau RB, Schoendorf A, Jennewein S;
XX
DR WPI: 2001-355489/37.
XX
DR P-PSDB; AAU02829.
XX
PT Novel nucleic acid and amino acid sequences, isolated from the Taxus
PT genus, useful for the synthetic production of Taxol and related
PT taxoids, intermediates within the Taxol biosynthetic pathway, and
PT other taxoid derivatives.
XX
XX
PS Claim 4; Page 101; 141pp; English.
XX
CC The sequence represents a DNA which encodes a Taxus cuspidata oxygenase.
CC The sequences isolated from the Taxus genus, and the respective
CC oxygenases are useful for the synthetic production of Taxol and related

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:23:04 ; Search time 1717.1 Seconds  
(without alignments) 9031.507 Million cell updates/sec

Title: US-09-995-917A-3  
Sequence: 1 atggagatgccttcattcg.....caaatcctgaagattag 1149

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estbda.\*  
2: em\_estbhm.\*  
3: em\_estlin.\*  
4: em\_estlmu.\*  
5: em\_estcov.\*  
6: em\_estcpl.\*  
7: em\_estcro.\*  
8: em\_estc.\*  
9: gb\_estl.\*  
10: gb\_estl2.\*  
11: gb\_hlc.\*  
12: gb\_gss.\*  
13: em\_gss\_hum.\*  
14: em\_gss\_inv.\*  
15: em\_gss\_pln.\*  
16: em\_gss\_vrl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237.4	20.2	755	10	BI926091 EST545980
2	206	17.9	574	9	AM615836 EST325334
3	204.4	17.8	574	9	AM615967 EST325333
4	181.8	15.8	615	9	AW398669 EST309169
5	164.6	14.3	609	10	BI204653 EST522683
6	163	14.2	678	10	BI204446 EST522683
7	157	13.7	533	10	BE450142 EST522683
8	153.2	13.2	667	10	BE450142 EST522683
9	143.4	12.3	621	10	BI204438 EST522683
10	128.6	11.2	522	10	BI204438 EST522683
11	127.2	11.1	573	10	BI206239 EST522683
12	124.8	10.9	660	10	BI266196 EST522683
13	120.2	10.5	582	10	BI205718 EST522683
14	118.2	10.3	642	10	BE315722 EST522683
15	116	10.1	533	10	BE315722 EST522683
16	110.4	9.6	477	10	BF050501 EST435659
17	108	9.4	543	10	BI432565 EST535326

18	98.4	8.6	586	10	BM409833	BM409833 EST584160
19	96.8	8.4	299	10	BI072383	BI072383 C074P58U
20	95	8.3	574	10	BI787249	BI787249 sag73b09.
21	93	8.1	312	10	D15214	D15214 RICC0279A.R
22	92.8	8.1	594	10	BI928055	BI928055 EST547944
23	91.4	8.0	536	9	AI994171	AI994171 701499941
24	89.2	7.9	524	9	AM690507	AM690507 NF035D01S
25	89.2	7.8	432	10	BI071486	BI071486 C058P32U
26	87.8	7.6	432	10	Z26124	Z26124 ATT51396.G
27	87.2	7.6	535	10	BI432249	BI432249 EST535010
28	86.6	7.5	788	10	BI970911	BI970911 CM830011B
29	86.4	7.5	684	10	Z29017	Z29017 ATT52030.G
30	85.4	7.5	684	10	Z29018	Z29018 ATT52031.G
31	83.8	7.3	1023	9	AI943419	AI943419 MF02A4.MF
32	83.6	7.3	343	10	T22324	T22324 4332.Lambda
33	82.6	7.2	608	10	BE124630	BE124630 EST393665
34	82.2	7.2	658	10	BI263183	BI263183 NF088D09P
35	82	7.1	360	10	T43151	T43151 6414.Lambda
36	81.8	7.1	501	10	T43286	T43286 6549.Lambda
37	81.6	7.1	619	10	BF424875	BF424875 SUS2D05.Y
38	81.6	7.1	760	12	BM479918	BM479918 BCGJ72TF
39	81.2	7.1	505	9	AM201483	AM201483 sF03f10.Y
40	80.8	7.0	461	9	AM759817	AM759817 s154d10.Y
41	80.8	7.0	508	10	BF595940	BF595940 sU67g12.Y
42	80.4	7.0	285	10	Z17988	Z17988 ATT50472.AC
43	80	7.0	469	9	AI993171	AI993171 701495585
44	80	7.0	686	9	BE039087	BE039087 AB09E01.A
45	79.6	6.9	339	10	T22325	T22325 4333.Lambda

#### ALIGNMENTS

RESULT 1  
BI926091  
LOCUS  
DEFINITION  
EST545980 tomato flower, buds 0-3 mm Lycopersicon esculentum CDNA  
clone CTOA28L21 5' end, mRNA sequence.  
ACCESSION  
BI926091  
VERSION  
BI926091.1 GI:16233588  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE  
AUTHORS  
1 (bases 1 to 755)  
Van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,  
Uterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,  
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)  
Unpublished (2001)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute

#### FEATURES

source

Seq primer: T3.  
Location/Qualifiers  
1..755  
/organism="Lycopersicon esculentum"  
/cultivar="RA496"  
/db\_xref="taxon:4081"  
/clone="CTOA28L21"  
/clone\_lib="tomato flower, buds 0-3 mm"  
/tissue\_type="flower"  
/dev\_stage="0-3mm buds"  
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:  
XhoI; supplier: Cornell University; sequencing: The  
Institute for Genomic Research; Flower buds and flowers

were taken from greenhouse plants (4-8 wks old, TA496).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen.

BASE COUNT 226 a 140 c 124 g 265 t  
ORIGIN

Query Match 20.2%; Score 232.4; DB 10; Length 755;  
Best Local Similarity 63.0%; Pred. No. 6.5e-46;  
Matches 359; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

2 tggagtgcccttcattgagaaactattcttcttcaaacctcatagatcagactcca 61  
Db 185 ttggatggccttacttgatggtaaaccttcttcttgaagcctcatcctttatgtca 244  
Qy 62 tcggtacattcttgaacaacagtggttcaagtgatgaaagtggttcaagtaatat 121  
Db 245 ttggatgcttcttcaacacatggtttagatgggaaagtggttcaagtcacatttat 304  
Qy 122 gtggtggaagcagtagtctcatgtgaccagaactcaactgttctacttcaaacg 181  
Db 305 tttttccccaacagtggtgacacacacacacacacacacacacacacacacacg 364  
Qy 182 aaggggaagtggttaccatggaattccaaagcagatgacatcttcgcgaatat 241  
Db 365 aagatgaattatttcagtgatgacacacacacacacacacacacacacacacacg 424  
Qy 242 ccccttattagccacagagaattccagagaactaaatgttattatagcttca 301  
Db 425 catgctgttgctgtggcgacacacacacacacacacacacacacacacacacg 484  
Qy 302 tcaatcccaagtggaac 361  
Db 485 tcagac 544  
Qy 362 tactaagcagtggaac 421  
Db 545 ttctccac 604  
Qy 422 ctctcagtgatgtgaacaacacacacacacacacacacacacacacacacacacac 481  
Db 605 catctcatgagatgagac 664  
Qy 482 atgtatgcaagatttttacttataagaggttatactccttaccacacacacacac 541  
Db 665 taattcttcaagatttttcttcttattatgagagatttaattcttaccattatgacac 724  
Qy 542 caggaaacggttatcaaacgcaattaag 571  
Db 725 ctggaac 754

RESULT 2  
AM615836 574 bp mRNA linear EST 18-MAY-2001  
LOCUS AM615836 tomato flower buds 0-3 mm, Cornell University  
DEFINITION Lycopersicon esculentum cDNA clone cTOA17E20 5', mRNA sequence.  
ACCESSION AM615836  
VERSION AM615836.1 GI:7321721  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.

REFERENCE 1 (bases 1 to 574)  
AUTHORS van der Hoeven, R.S., Bezzeredes, J.L., Matern, A.L., Holt, I.E., Liang  
F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M., Niernan, W.,  
Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.  
TITLE Generation of ESTs from tomato flower tissue, 0-3 mm buds  
JOURNAL Unpublished (1999)  
COMMENT Contact: CUGI

Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source Location/Qualifiers

1. 574  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cTOA17E20"  
/clone\_id="tomato flower buds 0-3 mm, Cornell University"  
/tissue\_type="flower"  
/dev\_stage="0-3mm buds"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Tanksley; Flower buds and flowers were  
taken from greenhouse plants (4-8 wks old, TA496). They  
were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."  
BASE COUNT 194 a 96 c 108 g 175 t 1 others  
ORIGIN

Query Match 17.9%; Score 206; DB 9; Length 574;  
Best Local Similarity 61.3%; Pred. No. 1.6e-35;  
Matches 332; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

Qy 75 ggaacaagtggttcaagtgatggaagaagtggtcaagtcacaatatatgtgtggaagaagc 134  
Db 1 gcaaac 60  
Qy 135 agtagctcattgtgac 194  
Db 61 agtgggtcattgtgac 120  
Qy 195 tcaatccgaattccaaagacagatgacacacacacacacacacacacacacacacacac 254  
Db 121 tcagtgatgattccaaagcacaattcattgacacacacacacacacacacacacacac 180  
Qy 255 caacggagaattccaaagac 314  
Db 181 tcttgccgac 240  
Qy 315 gtccgaac 374  
Db 241 gtcctaac 300  
Qy 375 gaaaaattgcggaagtggaattccataaagaagttaaatgttactcagtgat 434  
Db 301 gaaagatgaac 360  
Qy 435 ggtaaacgaactcttgagac 494  
Db 361 agtgaagac 420  
Qy 495 tttttatctataagaaggttatactccttaccacacacacacacacacacacacacac 554  
Db 421 tttttctgtcttgaagagatttaattcttccattatgacacacacacacacacacac 480  
Qy 555 tacaacgaacatgaagtgtagatccatcgtaatatatacaaaagcaattatagaaga 614  
Db 481 tgcagaagac 540  
Qy 615 ca 616  
Db 541 aa 542

RESULT 3  
AM615967 574 bp mRNA linear EST 18-MAY-2001  
LOCUS AM615967 tomato flower buds 0-3 mm, Cornell University  
DEFINITION Lycopersicon esculentum cDNA clone cTOA17E16 5', mRNA sequence.

```

ACCESSION   AM615967
VERSION     AM615967.1
KEYWORDS    GI:7321720
SOURCE      tomato.
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1 (bases 1 to 574)
AUTHORS    van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
            ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
            Frazer,C.M., Martin,G.B., Giovannoni,J.J., and Tanksley,S.D.
            Generation of ESTs from tomato flower tissue, 0-3 mm buds
            Unpublished (1999)
JOURNAL     Contact: CUGI
COMMENT     Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES
source      Location/Qualifiers
            1..574
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="CTOAL7E16"
            /clone_1lb="tomato flower buds 0-3 mm, Cornell University"
            /tissue_type="flower"
            /dev_stage="0-3mm buds"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; supplier: Tanksley; flower buds and flowers were
            taken from greenhouse plants (4-8 wks old, TA496). They
            were immediately frozen in liquid nitrogen and then
            size-separated while remaining frozen."
BASE COUNT  196 a          96 c          109 g          173 t
ORIGIN
Query Match      17.8%; Score 204.4; DB 9; Length 574;
Best Local Similarity 61.1%; Pred. No. 3.9e-39;
Matches 331; Conservative 0; Mismatches 211; Indels 0; Gaps 0;
QY 75 gcaacaagtgcttcacgtagaagaagtgctcaagaatatactgctggaagaagc 134
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1 GCACACACATTTGCTTACGATGGAAGTGTCAAGTACATTTATTTTCCCCCACC 60
QY 135 agtagctcatgtgacgaagaactcaactgttctactctcaaaagaagggaagttgt 194
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 61 AGTGTCATGTGACCAAGACCTTAATTACTTCATATTACAAAGCAAGATTAAGTTATT 120
QY 195 tcatcgtatataccaagaagcgtacatctcgcgaataatctccctctattagc 254
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 121 TCAGTGTACTTATCCAAAGCCCAATTCATGCTATGCGCAAAAGTTTCTGTGTGCG 180
QY 255 caacggagaatctcacaggaataaataaagtctattatagctctcatcaatcaca 314
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 181 TCTTGCGCGACACATAAAAGCCTTAGAATGTTTCATTATTCATACATACACACATTAA 240
QY 315 gtccgaacactgactctctcactgcgcagagaacactctctatctcgatcaagtcag 374
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 241 GCTTAACCTGAGTTTATTAATGATGTTGAAACATTAGCATTCCAGATTCTCCAAATCAG 300
QY 375 gaaaaattgcgagaagtcgaatctccataaagaagctaaagtcttactcagtcgtat 434
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 301 GAAAGATTAACATCAAGTCAGATACAGAGAGAGCAAGCAAGGATTTCCTTAATGTGAT 360
QY 435 ggtaaaccaacactctgagatcaagccagaaagaccagcaagaacttatgatctycaaga 494
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 361 AGTGAAGCAAGTACTTGGATTAACTCCAGATTAATCCACAAGCTGCAATTATCTTCAGCA 420
QY 495 tcttctatctatcagaaagggttatactcctaccgaataacgcgtctccaggaagsgtata 554
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

DB 421 TTTTCTGCTTTATAGAGGATTAAGTCTTTACCATTAATACATACCTGGAATCCATTA 480
QY 555 tacaacgaatgaagtagtagatcccaatcgtatatacatcaaacgaactatagaaga 614
    ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
DB 481 TCGAAGACGAGTCGACGCTAGAGATGATATCTTCACTATCAATCAATTAATAGCA 540
QY 615 ca 616
DB 541 AA 542
RESULT 4
LOCUS      AM398669
DEFINITION EST309169 L. pennellii trichome, Cornell University Lycopersicon
            pennellii cDNA clone cLPT4K22 5', mRNA sequence.
ACCESSION  AM398669
VERSION     AM398669
KEYWORDS    GI:6917139
SOURCE      EST.
ORGANISM    Lycopersicon pennellii.
            Lycopersicon pennellii.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1 (bases 1 to 615)
AUTHORS    Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
            Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii
            ,C.Y., Bowman,C.L., Nierman,W., Frazer,C.M., Venter,J.C., Martin
            ,G.B., Tanksley,S.D. and Giovannoni,J.
            Generation of ESTs from wild tomato (Lycopersicon pennellii)
            trichomes
            Unpublished (1999)
JOURNAL     Contact: CUGI
COMMENT     Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES
source      Location/Qualifiers
            1..615
            /organism="Lycopersicon pennellii"
            /db_xref="taxon:28526"
            /clone="cLPT4K22"
            /clone_1lb="L. pennellii trichome, Cornell University"
            /tissue_type="trichome"
            /dev_stage="mixed stages"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Leaves of various stages were shaken in liquid
            nitrogen, shearing off trichomes. This procedure yielded a
            mixture of cells highly enriched for trichomes, with minor
            contamination by other types of leaf cells."
BASE COUNT  180 a          111 c          100 g          224 t
ORIGIN
Query Match      15.8%; Score 181.8; DB 9; Length 615;
Best Local Similarity 63.7%; Pred. No. 1.2e-33;
Matches 276; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
QY 2 tgggaatgccttcatatggaagaactattctcttcaaacctcatagatcagactca 61
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 182 TTGGATGGGCTTTACTTGGTGAACCTTCTCTTTTGAAGCTCATCTCTTAATTCTTA 241
QY 62 tccgtacattcttgcacaacagtgcttcaacgtagaagaagtgctcaagtcataatcat 121
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 242 TTGGTACTTTCCTTCMAACATTTGTTAGGTATGGGAAAGTGTTCAGATCACATTAT 301
QY 122 gtggggaagaagcagtagtctcatgtgacgaagaactcaaatgtctactctcaaacg 181
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 302 TTTTCTCCCAACAGTGTGTCATGTGACCAAGACCTTAATTACTTAATTAACAAATG 361

```

```

QY 182 aagggaagttgttaccatcgatgattccaaaagcgaatgacattctcggcaaatatt 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 AAGATAGATTATTCAGTGTAGTATTCACAGCCAACTGATGTACTTGGCAAGCTTT 421
QY 242 ccccttataaccacccggaagaattcacaggaactaaanaattattatattagttca 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 CATGTGTTGGCTGTGGTGACACACATAAAGGCTTGAATGTTTCATTATCACTAA 481
QY 302 tcaatccacaaagtcgaacactgtacttctcactcgcgcagaagaacctctatctga 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 TCAGCACCATTTAAGTCTAAACCTGAGCTTATTATGATGTGTGAACATTAGCATTGACA 541
QY 362 tactaaagcatgaaaattccgcgaagtcgaattccataagaagttaaatttta 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 542 TTCTTTAAATCATGGAAGATTAACATCAAGTCAGATACCTGGGAGGAGCAAGAAATTTT 601
QY 422 ctctcagttgtat 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 602 CATCAATGTGAT 614

RESULT 5
BI204653 609 bp mRNA linear EST 11-JUL-2001
LOCUS EST522693 cTOS Lycopersicon esculentum cDNA clone cTOS5N9 5' end,
DEFINITION mRNA sequence.
ACCESSION BI204653
VERSION BI204653.1 GI:14682377
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 609)
AUTHORS Van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R.,
TITLE Ronning,C. and Tanksley,S.
JOURNAL Generation of ESTs from Tomato Suspension Cultures
COMMENT Unpublished (2001)
CONTACT: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
source
1. 609
/organism="Lycopersicon esculentum"
/cultivar="TA496, E6203"
/db_xref="taxon:4081"
/clone="cTOS5N9"
/tissue_type="suspension cultures"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Suspension cultures of L.esculentum E6203 were grown
in Murashige and Skoog based medium, supplemented with 15%
coconut milk (filter sterilized and added after
autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8).
Fresh medium was added every 7 days, and cultures were
grown at 25 C, with 12hrs of light and continuous
shaking."

BASE COUNT 209 a 101 c 124 g 175 t
ORIGIN
Query Match 14.3%; Score 164.6; DB 10; Length 609;
Best Local Similarity 64.5%; Pred.No.1.8e-29;
Matches 273; Conservative 0; Mismatches 129; Indels 21; Gaps 1;
QY 705 ggaatgggaacttttgaatgggaatgatacgaagaatgcaatgtgat 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 24 GGAAGAGAGCTTATGTGATTTGGAGATTTACCAAAAGATGGACTTCACTCAAAAGGTAAT 83

```

```

QY 765 ttctgagcactacagatgtggttaataatcgtaagactgtacataagaaaagctactcatga 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84 AANTGAAGCTTTAAGATATAGTATGGGAATGTTGTCAAAATTTGTGCACCGAAGCACTTAAGA 143
QY 825 tattaattcaagaatataatgattccaaagggtgaaagtgattccaatcttcacagc 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 144 TGTCAATATTAAAGATTTATGTATGATTCACCGGGTTGGAGAGTCATCAGTATTAAGTGC 203
QY 885 agtacaatcttgatccctctctcattcagaataatcctttgatttaattccatagatag-- 942
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 TGTTCATTTGGACCCATCAGTTCACTATGACATCCATTCCTTAATCTTGGAGATGGA 263
QY 943 -----acaaaagcagcggtttgaggagagtaaggtatgtcc 983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 GAGTGTATGACAAATATAGCAAGAAAGTTGACTCTTTTGGGGAGAGATCAAGATGTTGTC 323
QY 984 tgggtgtgaactgtgcaagctcccaattgcttctccttcatacatctgtcctctcta 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 TGGATTTGAACCTGCAAAAGTTGAACTGAGCTTCTTCTTCACACACCTGTACAAAAATA 383
QY 1044 taggtggaataaataatgcagatgaatgcaatgcacatcgcaccccttaagtgaagtaag 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 384 CAGATGGAGGTAGAGAGAGAGAACACCACTTGCTTATCCATATGTGAGATTCAAAAA 443
QY 1104 agg 1106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 444 TGG 446

RESULT 6
BI204446 678 bp mRNA linear EST 11-JUL-2001
LOCUS EST522486 cTOS Lycopersicon esculentum cDNA clone cTOS5E13 5' end,
DEFINITION mRNA sequence.
ACCESSION BI204446
VERSION BI204446.1 GI:14682170
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 678)
AUTHORS Van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R.,
TITLE Ronning,C. and Tanksley,S.
JOURNAL Generation of ESTs from Tomato Suspension Cultures
COMMENT Unpublished (2001)
CONTACT: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
source
1. 678
/organism="Lycopersicon esculentum"
/cultivar="TA496, E6203"
/db_xref="taxon:4081"
/clone="cTOS5E13"
/tissue_type="suspension cultures"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Suspension cultures of L.esculentum E6203 were grown
in Murashige and Skoog based medium, supplemented with 15%
coconut milk (filter sterilized and added after
autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8).
Fresh medium was added every 7 days, and cultures were
grown at 25 C, with 12hrs of light and continuous
shaking."

BASE COUNT 226 a 108 c 134 g 210 t
ORIGIN

```



High quality sequence stop: 420.

## FEATURES

source

Location/Qualifiers

1. .667  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-1781"  
/clone\_lib="Gm-cl066"  
/tissue\_type="Leaf and shoot tip, salt stressed, 2 week old seedling"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from unexpanded leaves and the shoot tips of 2 week old seedling from the cultivar Williams. The 2 week old seedlings were salt stressed in a solution of 500mM NaCl for 3 days prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 194 a 155 c 141 g 177 t  
ORIGIN

Query Match 13.2%; Score 152.2; DB 10; Length 667;  
Best Local Similarity 57.4%; Pred. No. 1.8e-26;

Matches 274; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

QY 1 atggatgagccttcatttgagagaactattcttctcaaacctatgacgtcc 60  
DB 128 ATGGGTTGGCCATTTCTTGGTGAACCATTTGGCTATTGAAGCTTATCTGCCACCA 187  
QY 61 atcgtacattcttgaacaacggtgttcacggtatggaagaagtgtcaacata 120  
DB 188 ATTAGGGGAATTCATGAGACACATAGCAAGTATGTCATTTACAAGTCAAACTG 247  
QY 121 tctgtggaagaacagtagtctcatgtgacaaagaactcaacatgtctactcaaac 180  
DB 248 TTTGGGGACCCAGCATAGTGTGACAGATGACGACTCAACAGGTTCACTTACAAC 307  
QY 181 gaaggagaattgttcatcgattatccaagaagcgtatgacattctcgtcaaat 240  
DB 308 GAAGGGAATTTTCAGTGCAGCTATCTAGAACATCGGTGGAATCTAGGAATAAG 367  
QY 241 tccctctatagccacgagaattcacagaagaactaaatgttattatgattc 300  
DB 368 TCCATGTGTGCTTAACTGTGATGATGATGATGATGATGATGATGATGATGATG 427  
QY 301 atcaatctacaagaagtcgaacgttcttctactgcgacagaacgtctctatc 360  
DB 428 CTAAAGCCAGCCAGCTCAGAACACACTCTTGAAGAGGTGAGAGATATCTCTTG 487  
QY 361 atactaagaatcagaagaattgacgaagtcgaattccataaagaatgaattt 420  
DB 488 GTTCTGAACCTCTTGGAGCCAAAGATTCATATTCCTGAGCCCAAGATGAAGTTC 547  
QY 421 actctagtgtagttaaaccacacttgagcatcaaggcaagaaccagaaga 477  
DB 548 ACCTTCAATGTAATGCTAAGCATATCATGAGATGATGATGATGATGATGATG 604

RESULT 9  
BI204438 621 bp mRNA linear EST 11-JUL-2001  
LOCUS BI204438  
DEFINITION EST522478 cTOS lycopersicon esculentum cDNA clone cTOS5C19 5' end,  
ACCESSION BI204438  
VERSION BI204438.1 GI:14682162

## KEYWORDS

EST.

tomato.

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 621)

van der Hoeven, R., Sun, H., Bezzerides, J., Cho, J., Utterback, R.,

Romling, C. and Tanksley, S.

Generation of ESTs from Tomato Suspension Cultures

Unpublished (2001)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

Location/Qualifiers

1. .621

/organism="Lycopersicon esculentum"

/cultivar="FA496, E6203"

/db\_xref="taxon:4081"

/clone="cTOS5C19"

/clone\_lib="cTOS"

/tissue\_type="suspension cultures"

/lab\_host="SOLR"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; Suspension cultures of L. esculentum E6203 were grown in Murashige and Skoog based medium, supplemented with 15% coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.0). Fresh medium was added every 7 days, and cultures were grown at 25 C, with 12hrs of light and continuous shaking."

## FEATURES

source

BASE COUNT 201 a 101 c 118 g 201 t  
ORIGIN

Query Match 12.5%; Score 143.4; DB 10; Length 621;  
Best Local Similarity 63.5%; Pred. No. 2.5e-24;

Matches 247; Conservative 0; Mismatches 121; Indels 21; Gaps 1;

QY 739 aagatgaattcaatcgaatgtgattcttgagcactacagatggttaataatcgtcaag 798  
DB 1 AAGATGACTTCACTCAAAAAGTAATAAATGAAGCTTTAAGATATGGAATGTTGCAAA 60  
QY 799 actgtacataagaagaactcatcatgatataattcaagaagaatgtgattccaaagg 858  
DB 61 TTTGTCCACCGAAGACCTTTAAAGATGTCAAAATTTAAAGATTATGTGATCCAGCGG 120  
QY 859 ttggaagtggttccaatcttcaacagcagtcacatcgtgatccctctctcatgaata 918  
DB 121 TGGAAAGTCTTACCATATTTCAAGTGTCTGTTCAATTTGGACCCATCACTTCTAATG 180  
QY 919 ttggaatttaatlccatlgagatgg-----accaaacgacgagcg 957  
DB 181 CTCACCTTAAATCTTGGAGATGAGAGATGATGATGATGATGATGATGATGATGATG 240  
QY 958 ttggaaggaaggaatgaaggtatgtctgtgtggaacttggcgaagtcgaatgttct 1017  
DB 241 TTTGGGGAGAGATCAAGATGTTGCTCTGATTTGAATTTGCAAGTGTGAAGTACCTTC 300  
QY 1018 ttcttcatacatcttgcctctctataggtggaataataaagtcagatgaatgccaatc 1077  
DB 301 TTCTTTCACCACTTGTACAAAATATACAGATGGAGGTAGAGAGAGAGACCAACCAT 360  
QY 1078 gcgcacccttaagtgagtttaagagag 1106  
DB 361 GCTTATCATATGTGGAGTCAAAATAG 389

RESULT 10  
BG316131







<b>REFERENCE</b>	1 (bases 1 to 582)
<b>AUTHORS</b>	van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R., Romling,C. and Tanksley,S.
<b>TITLE</b>	Generation of ESTs from Tomato Suspension Cultures
<b>JOURNAL</b>	Unpublished (2001)
<b>COMMENT</b>	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> . Location/Qualifiers
<b>FEATURES</b>	1..582
<b>SOURCE</b>	/organism="Lycopersicon esculentum" /cultivar="TA496, E6203" /db_xref="taxon:4081" /clone="CTOS8H16" /clone_1lb="CTOS" /tissue_type="Suspension cultures" /lab_host="SOLR" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Suspension cultures of L.esculentum E6203 were grown in Murashige and Skoog based medium, supplemented with 15% coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8). Fresh medium was added every 7 days, and cultures were grown at 25 C, with 12hrs of light and continuous shaking."
<b>BASE COUNT</b>	182 a 111 g 192 t
<b>ORIGIN</b>	
Query Match	10.5%; Score 120.2; DB 10; Length 582;
Best Local Similarity	62.6%; Pred. No. 1e-18;
Matches 216; Conservative 0; Mismatches 108; Indels 21; Gaps 1.	
Oy 783	tgtgaatatcgtcaagaactgctacataagaagaagctcctcatgtatataattcaagaata 842
Db 1	TGGGAATGTTGTCAAAATTGTGCACCGAAGGCACCTTAAAGTGCCAATTAAAGAATTAA 60
Oy 843	tgtgatcccaaagggttgaaaggctgtcccaattccacaagtagtacatctgacctc 902
Db 61	TGTGATTCACACGGGTTGGAAGAGTCCTACACCATATTCAGTGTCTTCATTGTGGACCACATC 120
Oy 903	tcttcataaaaaacccttttgaaattbaalccaatgagaatgg----- 942
Db 121	AGTTCATCTCTATGACACTCCACTTTAATTCCTTGAGATGTGGAGTAGATGACCAATAAG 180
Oy 943	-accaaaacgcgcgcttttgagaagaggttaaggatagtctcgctggttgaaactggcaa 1001
Db 181	CAGAAGATTGACTCTTTTGGGGGAGATCAAGAATGTTGTCTCGAATTTGAACCTGCCAAA 240
Oy 1002	gtcccaaatctcttctcttcatacacatctgctcctccctcctaagtaggaanaaagtc 1061
Db 241	GCTTGAACTTACCTCTTCTTCACCACTCTGTACAAAAAATACACATGAGGAGGTGAGAAGA 300
Oy 1062	agatgtaaattgccaatcgcgcaccccttaccgtgagattaaagagag 1106
Db 301	AGGAGAACCAACCCCATTTGCTTATCCATATGTGAGAGTTCAAATAATGG 345
<b>RESULT 14</b>	
BEJ15722	
LOCUS	642 bp mRNA linear EST 14-JUL-2000
DEFINITION	NF025011.F1090 Developing leaf Medicago truncatula cDNA clone
ACCESSION	NF025011.F1090
VERSION	BEJ15722
KEYWORDS	EST.
SOURCE	barrel medic. Medicago truncatula
ORGANISM	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoidae; Trifolieae;

[illegible]

SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 533)  
AUTHORS Alcala,J., Vrehalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
Niernan,W., Frazer,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,  
S.D.  
TITLE Generation of ESTs from tomato fruit tissue, immature green  
JOURNAL Unpublished (2000)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
Location/Qualifiers  
FEATURES  
source  
1..533  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEM18G22"  
/clone\_1fb="tomato developing/immature green fruit"  
/tissue\_type="fruit"  
/dev\_stage="immature green (5-35 days post-anthesis)"  
/lab\_host="SOLR"  
/note="Vector: pBluescriptSkmCudapt; Site\_1: EcoRI;  
Site\_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and  
harvested at 7 day intervals through 35 dpa. Equal masses  
of tissue from each stage were combined (including seeds  
and locules) prior to mRNA isolation."  
BASE COUNT 172 a 75 c 111 g 175 t  
ORIGIN  
Query Match 10.1%; Score 116; DB 10; Length 533;  
Best Local Similarity 56.6%; Pred No. 1,1e-17;  
Matches 215; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 1 atggatggccttcattgagaaactattcttcttccaacccaatagatcagactcc 60  
DB 148 ATGGGTTGGCCTTCTTCTGCTGAACCTATTGTTATTTGAGACCTATTACGCTACTACT 207  
QY 61 atcggtacattcttgcacaacacggtgttcaagtgatgaaaagtggttcaagtcacaata 120  
DB 208 ATTGGAGATTTCATGCACAGATCATATTCTAGGTATGGAGGAAATTTTCAAGTCAATTTTG 267  
QY 121 tgtgtgaaaagcagtagctcatgtgcaagaactcaacatggttcaacttcaaacaaac 180  
DB 268 TTTGGAGAGCCACACATATTCTTCAAGAGATGCAAGGCTTAACAGATACATTCTCGCAAA 327  
QY 181 gaagggaagtggttaccatcgatataccaaaagcgatgcacatcttcgcgcaaatat 240  
DB 328 GAAGGAGATTTATGATTAATTCACAGAGATAGTGGGATCTGTAATAGG 387  
QY 241 tccctctattgcccacgcgagagaatcacaagaaactaaaatggttattatagcttc 300  
DB 388 TCTATGTAGTTCAAGTGTGACAAATGACATAGATATAGAGATGATTCCTCGAATTTT 447  
QY 301 atcaatctcaaaaagtcgaacctgacttcttcaactgcgacagagaacctctatctcg 360  
DB 448 TTGAGCAATGCTAGGCTTAAGATCAACTTTTAAGTGAAGTTGAAAAGCATACTTGGCTT 507  
QY 361 atactaaagtcataggaaaaa 380  
DB 508 GTTCTTGGCTCTTGGAACA 527

Tue Jul 30 08:40:09 2002

us-09-995-917a-3.rst

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:28:34 ; Search time 50.3 Seconds  
(without alignments)

5610.992 Million cell updates/sec

Title: US-09-995-917A-3

Perfect score: 1149  
Sequence: 1 atggagtgcccttcattg.....caaatctctgaagattg 1149

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_NA:\*  
2: /cgn2\_6/prodata/2/1na/5A\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/1na/5B\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/1na/5A\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/1na/5B\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	97.8	8.5	1608	2	US-08-622-166A-1
2	80	7.0	4937	2	US-08-622-166A-3
3	58.6	5.1	7218	1	US-08-232-463-14
4	38.8	3.4	2694	1	US-08-465-995A-1
5	38.8	3.4	2694	2	US-08-465-994C-1
6	38.8	3.4	2694	2	US-08-966-145-1
7	38.8	3.4	2760	1	US-08-101-593-1
8	37.4	3.3	3892	4	US-08-569-214-1
9	37.4	3.3	3892	4	US-08-937-236-1
10	37.2	3.2	2694	1	US-08-465-995A-3
11	37.2	3.2	2694	2	US-08-465-994C-3
12	37.2	3.2	2694	2	US-08-966-145-3
13	37.2	3.2	2760	1	US-08-101-593-3
14	37.2	3.2	9636	1	US-08-323-170B-1
15	37	3.2	9636	4	US-08-954-441-1
16	36.8	3.2	3308	3	US-08-714-918-68
17	36.8	3.2	3308	4	US-09-265-315-68
18	36.8	3.2	3308	4	US-09-265-315-68
19	36.8	3.2	3308	4	US-09-266-417-68
20	36.2	3.2	1850	3	US-08-724-466B-3
21	36.2	3.2	1850	4	US-08-882-164D-3
22	36.2	3.2	29793	4	US-09-302-812-38
23	36.2	3.2	29793	4	US-09-311-477-38
24	35.8	3.1	4090	4	US-08-569-214-4
25	35.8	3.1	4090	4	US-08-937-236-4
26	35.2	3.1	1725	4	US-08-882-164D-31
27	35.2	3.1	2439	4	US-09-386-493-1

28	34.6	3.0	1659	1	US-08-231-729B-1	Sequence 1, Appl1
29	34.4	3.0	2277	1	US-08-676-967-2	Sequence 2, Appl1
30	34.4	3.0	2277	1	US-08-676-974-2	Sequence 2, Appl1
31	34.4	3.0	2277	2	US-09-098-487-2	Sequence 2, Appl1
32	34	3.0	1659	1	US-08-231-729B-2	Sequence 2, Appl1
33	33.6	2.9	1333	1	US-08-684-862-9	Sequence 9, Appl1
34	33.2	2.9	2659	3	US-09-029-267-19	Sequence 19, Appl1
35	33.2	2.9	8157	1	US-08-358-160-65	Sequence 65, Appl1
36	33.2	2.9	8584	1	US-08-358-160-66	Sequence 66, Appl1
37	33.2	2.9	8590	1	US-08-358-160-70	Sequence 70, Appl1
38	33	2.9	398	4	US-09-385-982-273	Sequence 273, App
39	33	2.9	2256	1	US-08-181-271A-2	Sequence 2, Appl1
40	33	2.9	2256	1	US-08-449-315-2	Sequence 2, Appl1
41	33	2.9	2256	1	US-08-444-803-2	Sequence 2, Appl1
42	33	2.9	2256	1	US-08-449-043-2	Sequence 2, Appl1
43	33	2.9	2256	1	US-08-456-265A-2	Sequence 2, Appl1
44	33	2.9	2256	1	US-08-455-416-2	Sequence 2, Appl1
45	33	2.9	2256	1	US-08-455-244-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-622-166A-1  
Sequence 1, Application US/08622166A  
Patent No. 5952545  
GENERAL INFORMATION:  
APPLICANT: KONCZ, CSABA  
APPLICANT: MATHUR, JAIDEEP  
APPLICANT: SZEKERES, MIKLOS  
APPLICANT: ALTMANN, THOMAS  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME  
TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/622,166A  
FILING DATE: 27-MAR-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SVENSSON, LEONARD R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 0147-0153P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1608 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: A. thaliana  
IMMEDIATE SOURCE:  
LIBRARY: lambda gt10  
CLONE: C204

```

;      FEATURE:
;      NAME/KEY:      CDS
;      LOCATION:      48..1466
US-08-622-166A-1

```

Query Match	8.5%;	Score 97.8;	DB 2;	Length 1608;
Best Local Similarity	58.9%;	Pred. No. 1,1e-18;		
Matches 168; Conservative	0;	Mismatches 117;	Indels 0;	Gaps 0

Oy	671	a t g a a g a c a t c y g c g c a t t a g a c c a a g a a g g a t b g g a c t t t g a a t t g g a a g	730
Db	961	A G C A A G A G C A T G A A A A T T T A G G S C A A T G A A G T A T T G T A T A C T C T T G A A T G A G A G	1026
Oy	731	a t t a t c a g a a g a t y g a a t t c a c t c a a t b g t y g a t t t c t y a a g g a c a g a t y t g t a t a	790
Db	1021	A T T C A A G T C A T A T C C A T T T A C A C A A A G T G T G T T A T A G A G C G T C A C A G T G G C T A A C	1086
Oy	791	t c g t a a g a c t g t a c t a g a a a g c t a c t a t a t a t t a a t c a a g a n a t a t g y a t t c	850
Db	1081	T C A T C G C G C G T T T T C A G A C G T C C A A T A G A C G A T G T T G A T A C C A A A G T T A T A A A A T T C	1146
Oy	851	c a a a g y g t g a a g t g t t t c a a t c t t c a c a g a t a c a t t i g a t t c c t c t t c a t y	910
Db	1141	C A A A A G G T G A A A G T A T T C T C A T C A C G T T T A G A C G G T T C A T T A T T A G A C C A A A C C A C T T C A	1206
Oy	911	a a a t c c t t t g a a t t a a t c c a t g a t a g a c a g a c a a a a g a c a g	955
Db	1201	A A G A T G C T C G C A C T T T C A A C C C T T T G G A G A T G G C A G A C A C A C T C G G	1245

RESULT 2  
US-08-622-166A-3

```

1  GENERAL INFORMATION:
2  APPLICANT: KONCZ, CSABA
3  APPLICANT: MATTHUR, JAIDEEP
4  APPLICANT: SZEKERES, MIKLOS
5  APPLICANT: ALTMANN, THOMAS
6  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
7  TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS I
8  NUMBER OF SEQUENCES: 4

```

STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/622,166A  
FILING DATE: 27-MAR-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SVENSSON, LEONARD R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 0147-0153P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4937 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

```

1 MOLECULE TYPE: DNA (genomic)
2
3 HYPOTHETICAL: NO
4
5 ANTI-SENSE: NO
6
7 ORIGINAL SOURCE:
8 ORGANISM: Arabidopsis thaliana
9 STRAIN: cv. Columbia
10 IMMEDIATE SOURCE:
11 LIBRARY: lambda gt10
12
13 CLONE: C204
14
15 FEATURE:
16
17 NAME/KEY: CDS
18 LOCATION: join(968..1483, 1660..1829, 1917..2165, 3903
19 LOCATION: ..3989, 4084..4162, 4248..4354, 4446..4576, 4674
20 LOCATION: ..4773)
21
22 US-08-622-166A-3

```

```

Query Match      7.0%; Score 80; DB 2; Length 4937;
Best Local Similarity 53.5%; Pred. No. 2.3e-13;
Matches 167; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

Oy	10	cccttcattggaagaactatctcttcctcaaacctcagaactcagcattcgttaca	69
Oy	1082	cctctgattagagagactttttcagctgagtcggagctttacaaaacagaaacctgagcct	1141
Oy	70	tctcttgcacaacagtgctttcaacggtatgagaaaagtgcttaagtcacatatatggttga	129
Db	1142	tttcattcgcgcgagaaagatgagcccgctacggttcgggttttcatacagcatgtttttgtata	1201
Oy	130	aaagcagtagtgcattcattgagcacaagaactcaacatgtttcatacttcaaaaacgaaaggaag	189
Db	1202	ccgacgattttttctcagctgaccccgaaacgaacccggtttgtttcttcagacaaagaggaag	1261
Oy	190	ttgtttacatctggaattatccaaaagcagatcagacattctcggaaatctcccttca	249
Db	1262	cttttttgagttcttttttccctgcttccattttgaaaccttttgggcaaaacactcttgcctt	1321
Oy	250	ttagccaccggaagaatcaccagagaactcaaaaatgattatattatagcttcaacatc	309
Db	1322	ctttatgaaaggtcttttgcataaacgatatcactctccacatgagctttgctaattct	1381
Oy	310	acaaagtcgaaa	321
Db	1382	ttcaatcattttaa	1393

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: ASCII DOS/TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465, 994C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MARGARET A. CHURCHILL
; REGISTRATION NUMBER: 39,944
; REFERENCE/DOCKET NUMBER: 1920-305D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213/977-1001
; TELEFAX: 213/977-1003
; INFORMATION FOR SEQ ID NO: 1:
; SROUENCE CHARACTERISTICS:
; LENGTH: 2694 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; LOCATION: 1..2694
; US-08-465-994C-1
```

```

Query Match          3.4%; Score 38.8; DB 2; Length 2694;
Best Local Similarity 45.2%; Pred. No. 0.12;
Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
```

```

QY 599 acccaattatagaagacatgaataatgaatgaagaagaatttcctgattcgataa 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 ATGAATATATCGAGAGACAGCTCCTAAACCAAGTATGTAATTTCTTGTTCCTCGAATG 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 ttccgaatgaagatgaagaacatgcagccattagaagccaaagaaggatgggaactt 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 GATGGATGATGATGAATGAACATCAGAAAGATGATCATTCCTCAAGAAATGCGTAAAGTATTTT 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 tgattgggaagatttcgaagaatgaatgaatcctcaatgtgtgattcttggaagcactac 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 TCCAGCGTAAAGATGTAAGAAAGAAATGTCGCTGAAGAAATGCGGAAGCTATTTA 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 779 gatgtgtaataatcgcgaagactgtacatagaagaagctactcatgtatattaatccaag 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 AAAAGATTTATTAAGAAAGCGCGAGGCTTTGTTCACATAACCAAGATTGAACGATATG 1549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 839 aatatgtgattccaaagggtggaagggtttccaatcttcacagcagatcattgac 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1550 TTAAAGTTCACTGATGATTTCTTAAATGAAGACTATGAAATTATGTAATCTGTTCTTAATA 1609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 cctctctcatgaa 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 GTCTGATGAAGAA 1623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 6
US-08-966-145-1
; Sequence 1, Application US/08966145A
; Patent No. 5945312
; GENERAL INFORMATION:
; APPLICANT: Goodman, Myron F.
; APPLICANT: Reha-Krantz, Linda J.
; TITLE OF INVENTION: Synthesis of Fluorophore-Labeled DNA
; FILE REFERENCE: 1920-353D1
; CURRENT APPLICATION NUMBER: US/08/966,145A
; CURRENT FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: US 08/632,742
; EARLIER FILING DATE: 1996-04-15
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 2694 base pairs
```

```

; TYPE: DNA
; ORGANISM: T2 bacteriophage
; PUBLICATION INFORMATION:
; AUTHORS: Goodman, Myron F.
; PATENT DOCUMENT NUMBER: US 5,660,980
; PATENT FILING DATE: 1995-06-06
; PUBLICATION DATE: 1997-08-26
; US-08-966-145-1
```

```

Query Match          3.4%; Score 38.8; DB 2; Length 2694;
Best Local Similarity 45.2%; Pred. No. 0.12;
Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
```

```

QY 599 acccaattatagaagacatgaataatgaatgaagaagaatttcctgattcgataa 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 atgaatatatcgcaggaacagctccctaaacaaagtatgataattcttcttcgaatg 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 ttccgaatgaagatgaagaacatgcagccattagaagccaaagaaggatgggaactt 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 gatggaatgaatgaagaacatgcagccattagaagccaaagaaggatgggaactt 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 tgattgggaagatttcgaagaatgaatgaatcctcaatgtgtgattcttggaagcactac 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 tccagcgtaagattggaagaaggatgctgttcttcaactaaccaagaatttgaagcattg 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 779 gatgtgtaataatcgcgaagactgtacatagaagaagctactcatgtatattaatccaag 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 aaaagattatagaagaaggcagggctgttcttcaactaaccaagaatttgaagcattg 1549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 839 aatatgtgattccaaagggtggaagggtttccaatcttcacagcagatcattgac 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1550 ttaagttcaatgtattcttcaatgaagaactatcgaattatcattcttcttaata 1609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 cctctctcatgaa 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 gtctgattgaaga 1623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 7
US-08-101-593-1
; Sequence 1, Application US/08101593
; Patent No. 5547859
; GENERAL INFORMATION:
; APPLICANT: Goodman, Myron F.
; APPLICANT: Reha-Krantz, Linda J.
; TITLE OF INVENTION: NEW DNA SEQUENCING ENZYMES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Robbins, Berliner & Carspm
; STREET: 201 No. 5547859th Figueroa Street, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90012-2628
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/101,593
; FILING DATE: 19930802
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Spitals, John P.
; REGISTRATION NUMBER: 29,215
; REFERENCE/DOCKET NUMBER: 1920-305
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 1:
```



```

: SEQUENCE CHARACTERISTICS:
:   LENGTH: 2760 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: DNA (genomic)
:   FEATURE:
:     NAME/KEY: CDS
:     LOCATION: 1..2760
:   US-08-101-593-1

```

```

Query Match          3.48; Score 38.8; DB 1; Length 2760;
Best Local Similarity 45.2%; Pred. No. 0.12;
Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

```

```

QY 599 acgcacatataagagacatgaataatgcaataagagaagaagatttctgattcgataa 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1376 ATGAATATATTCGACGAGAACAGCTCTTAACCAAGTATGATATCTTCTCGAATG 1435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 ttctgaatgaagatgaagaacatgcaagcattagagcagaagaaggatggggaacttt 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1436 GATGATGATGATGATGAACATCAAGAACGATATTCATTCAGAAAGAAATGCTAAAGTATTT 1495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 tgaattggagaattatcagaagaatggaattcaactcaatgtgtgattcttgaggcaactac 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1496 TCCAGCCTTAAGATGGAAGAAAGAAATGTTCCGTGAAGAAATGAATGCCAGACTATTA 1555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 779 gattgtgtaacatcgtcaagaactgcatagaaagactactatgattatgaattcaag 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1556 AAAAGATTATTAAGAAAGCGCGAGGCTGTCTCACTAACAACCAAGATGGAACGATATG 1615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 839 aatatggaatccaaaggagggtggaagggtgttccaaatcttcacagcagatcattgac 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1616 TTAAGTTCACTGATGATTTCTTAAATGAATGATGATGATGATGATGATGATGATGAT 1675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 cctctctcatgaa 912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1676 GTCGTATGTAAGAA 1689

```

```

RESULT 8
US-08-569-214-1
: Sequence 1, Application US/08569214
: Patent No. 6155469
: GENERAL INFORMATION:
:   APPLICANT: MANN, BARBARA J.
:   TITLE OF INVENTION: RECOMBINANT ENTAMOEBA HISTOLYTICA LECTIN
:   TITLE OF INVENTION: SUBUNIT PEPTIDES AND REAGENTS SPECIFIC FOR MEMBERS OF THE
:   NUMBER OF INVENTION: 170 KD SUBUNIT MULTIGENE FAMILY
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: MORRISON & FOERSTER
:     STREET: 2000 PENNSYLVANIA AVENUE N.W., STE. 5500
:     CITY: WASHINGTON
:     STATE: DC
:     COUNTRY: USA
:     ZIP: 200006-1812
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentin Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/569,214
:     FILING DATE:
:   CLASSIFICATION: 424
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: PCT/US94/06890
:     FILING DATE: 17-JUN-1994
:     ATTORNEY/AGENT INFORMATION:
:       NAME: MURASHIGE, KATE H.

```

```

: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 9148-0006.21
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (202) 887-1500
:   TELEFAX: (202) 887-0763
:   TELEX: 90-4030
:   INFORMATION FOR SEQ ID NO: 1:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 3892 base pairs
:       TYPE: nucleic acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:     FEATURE:
:       NAME/KEY: CDS
:       LOCATION: join(1..3873, 3877..3882, 3886..3891)
:     US-08-569-214-1

```

```

Query Match          3.3%; Score 37.4; DB 4; Length 3892;
Best Local Similarity 53.8%; Pred. No. 0.35;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```

```

QY 573 tagatcaatcgttaatacatcaaaacgcaattagagaagcatgaaatgcaataag 632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 TAAATATATCAATGAAGATCATATATGTAAGCTTGACATATTAATAAACAATAAT 293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 633 agaagaagatttctggaatcgaataatcgaatgaatgaagaacatgagccatag 692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TAAACAAGATTTTGTCAAAAAGAAATATGATATCAATGTAAGAAATGATGATGATG 353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 693 agccaagaaggagtgagggaac 715
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 GGACAAATGTTCCAGTTGATGAAC 376

```

```

RESULT 9
US-08-937-236-1
: Sequence 1, Application US/08937236
: Patent No. 6187310
: GENERAL INFORMATION:
:   APPLICANT: MANN, BARBARA J.
:   APPLICANT: PETRI, WILLIAM A.
:   TITLE OF INVENTION: RECOMBINANT ENTAMOEBA HISTOLYTICA LECTIN
:   TITLE OF INVENTION: SUBUNIT PEPTIDES AND REAGENTS SPECIFIC FOR MEMBERS OF THE
:   NUMBER OF INVENTION: 170 KD SUBUNIT MULTIGENE FAMILY
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: MORRISON & FOERSTER
:     STREET: 2000 PENNSYLVANIA AVENUE N.W., STE. 5500
:     CITY: WASHINGTON
:     STATE: DC
:     COUNTRY: USA
:     ZIP: 200006-1812
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: Patentin Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/937,236
:     FILING DATE:
:   CLASSIFICATION: 424
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: 08/569,214
:     FILING DATE: 16 SEPTEMBER 1997
:     ATTORNEY/AGENT INFORMATION:
:       NAME: LIYNAT, SHMUEL
:       REGISTRATION NUMBER: 33,949
:       REFERENCE/DOCKET NUMBER: 291482000622
:       TELECOMMUNICATION INFORMATION:
:         TELEPHONE: (202) 887-1500
:         TELEFAX: (202) 887-0763

```

TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3892 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..3873, 3877..3882, 3886..3891)  
; US-08-937-236-1

Query Match 3.3%; Score 37.4; DB 4; Length 3892;  
Best Local Similarity 53.8%; Pred. No. 0.35;  
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
QY 573 tagatccaatcgtaatacatcaaaaagcaattatagaagacatgaataatgcataag 632  
DB 234 TAAATACTATCAATCAAGAGATCATATGTAAAGTTGAAGATTTAAATAAACCAATAT 293  
QY 633 agagaagaattttcgtgattcgtaattcgaaatgaagaatgaagaacatgcacatag 692  
DB 294 TAAACAAGATTTTTCGCAAAAAGATATGCAATTCATTCGAATTTGAAAATATGAAGTTGAT 353  
QY 693 agccaagaaggggagatggggaac 715  
DB 354 GGACAAATGTTCCAGTTGATGAAC 376

RESULT 10  
US-08-465-995A-3  
; Sequence 3, Application US/08465995A  
; Patent No. 5660980  
; GENERAL INFORMATION:  
; APPLICANT: Myron F. Goodman  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING AND ISOLATING  
; TITLE OF INVENTION: VARIANT T4 POLYMERASES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson, LLP  
; STREET: 201 No. 5660980th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90012-2628  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: ASCII DOS/TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465, 995A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Robert Berliner  
; REGISTRATION NUMBER: 20,121  
; REFERENCE/DOCKET NUMBER: 1920-305D2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 213/977-1001  
; TELEFAX: 213/977-1003  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2694 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2694  
; US-08-465-995A-3

Query Match 3.2%; Score 37.2; DB 1; Length 2694;  
Best Local Similarity 44.9%; Pred. No. 0.34;  
Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 agcaaatatagaagacatgaataatgaataagaagaagaattttcgtatcgataa 658  
DB 1310 ATGATATATTCGCAAGAACACTCTTAACCGAGATGATGATATTTCTGTCCTCGAATG 1369  
QY 659 ttccgaatgaagaatgaacatgcagccattagaagccaagaaggagatggggaattc 718  
DB 1370 GATGATATGATGAACATCATCAAGAGGTATCATTCCAAGGAAGATGCTAAAGTATTTT 1429  
QY 719 tgaattgggaagatataagaagaatggaattcactcaatggtgatttcgaagcactac 778  
DB 1430 TCCACGCTTAAGACTCGAAAAAATGTTGCTGGAAGAAATGAAATGCCGAAGCTATTA 1489  
QY 779 gatggtgaataatcgtaagaactgtacataagaagaagctactcatgataatcaatcaag 838  
DB 1490 AAAAGATTATTTGAAAGCCGACGCTTGTTCACATAAACCAAGAGTTGAACGATATG 1549  
QY 839 aatatgtatccaaagggtgggaagtggttccaattccaagcaggtacatctgac 898  
DB 1550 TTAAGTTCAGTATGATTTCTTAATGAACATGCAATTAACCGAATCTGTTCAATA 1609  
QY 899 cctcttcacatga 912  
DB 1610 GTCTGATTGAAGA 1623

RESULT 11  
US-08-465-994C-3  
; Sequence 3, Application US/08465994C  
; Patent No. 5928919  
; GENERAL INFORMATION:  
; APPLICANT: MYRON F. GOODMAN  
; TITLE OF INVENTION: LINDA L. REHA-KRANTZ  
; TITLE OF INVENTION: VARIANT DNA POLYMERASES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson, LLP  
; STREET: 201 No. 5928919th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90012-2628  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: ASCII DOS/TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465, 994C  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MARGARET A. CHURCHILL  
; REGISTRATION NUMBER: 39,944  
; REFERENCE/DOCKET NUMBER: 1920-305D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 213/977-1001  
; TELEFAX: 213/977-1003  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2694 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2694  
; US-08-465-994C-3

Query Match 3.2%; Score 37.2; DB 2; Length 2694;  
Best Local Similarity 44.9%; Pred. No. 0.34;  
Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagagaagaatttctgatactgataa 658  
DB 1310 ATGAAATATATCGACGAAACAGCTCTTAACCGAGTATGTAATCTGTGTTCCGAATG 1369  
QY 659 ttccgaattgaagatgaagacacgcagccatttgagcgaagaaggatgggaactt 718  
DB 1370 GATGAGTATGATTAACATCAAGAAAGTATCTTCCAAAGAAATCGCTAAAGTATTTT 1429  
QY 719 tgaattgggaagatltacgaagaatgaattcctcaattgtgattcttgaaggcactac 778  
DB 1430 TCCAGCGTAAAGACTGGAATAAATAATGTTCCTGAAGAAATGATGCCGAAGCTATTTA 1489  
QY 779 gatgtggaattatctgcacaagaactgtacatagaagaagctactcattatcaaatcaag 838  
DB 1490 AAAAGATTATTAAGAAAGCGGCTCTTGTCAACTTAACGAGAAATGTAACGATATG 1549  
QY 839 aatactgattccaaagggtgggaagggtttcccaattcttcacagcagtaacttgcac 898  
DB 1550 TTAAGTTCACTGATGATTTCTTAATGAACATATCGAATTACCGAATCTGTTCTCAATA 1609  
QY 899 cctctctcatgaa 912  
DB 1610 GTCTGATTGAAGAA 1623

## RESULT 12

US-08-966-145-3  
; Sequence 3, Application US/08966145A  
; Patent No. 5945312  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Myron F.  
; APPLICANT: Reha-Krantz, Linda J.  
; TITLE OF INVENTION: Synthesis of Fluorophore-labeled DNA  
; FILE REFERENCE: 1920-353D1  
; CURRENT APPLICATION NUMBER: US/08/966,145A  
; EARLIER FILING DATE: 1997-11-07  
; EARLIER APPLICATION NUMBER: US 08/632,742  
; NUMBER OF SEQ ID NOS: 4  
; SEQ ID NO 3  
; LENGTH: 2694 base pairs  
; TYPE: DNA  
; ORGANISM: T4 bacteriophage  
; PUBLICATION INFORMATION:  
; AUTHORS: Goodman, Myron F.  
; PATENT DOCUMENT NUMBER: US 5,660,980  
; PATENT FILING DATE: 1995-06-06  
; PUBLICATION DATE: 1997-08-26  
US-08-966-145-3

Query Match 3.2%; Score 37.2; DB 2; Length 2694;  
Best Local Similarity 44.9%; Pred. No. 0.34;  
Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagagaagaatttctgatactgataa 658  
DB 1310 atgaattatcgacgaagacagctcctaaaccgagtgaataattcttctgttcgcgaattg 1369  
QY 659 ttccgaattgaagatgaagacacgcagccatttgagcgaagaaggatgggaactt 718  
DB 1370 gatgtggaattatctgcacaagaactgtacatagaagaagctactcattatcaaatcaag 838  
QY 719 tgaattgggaagatltacgaagaatgaattcctcaattgtgattcttgaaggcactac 778  
DB 1430 tccagcgtaaagactcgaagaaagaaatcgtctcgaagaagaatgaatcgcgaagctatta 1489

QY 779 gatgtgtaataatctgcacaagaactgtacacalagaagaagctactcattgataatcaaaag 838  
DB 1490 aaaagaattatltagaaggcgagggcttcttcaactaaacagaagcttgaagatg 1549  
QY 839 aatactgattccaaagggtgggaagggtttcccaattcttcacagcagtaacttgcac 898  
DB 1550 ttaagtcgattgattcttcaaatgaactacgaatcaacacgaagctcgttcccaata 1609  
QY 899 cctctctcatgaa 912  
DB 1610 gctcgtattgaagaa 1623

## RESULT 13

US-08-101-593-3  
; Sequence 3, Application US/08101593  
; Patent No. 5547859  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Myron F.  
; APPLICANT: Reha-Krantz, Linda J.  
; TITLE OF INVENTION: NEW DNA SEQUENCING ENZYMES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carpm  
; STREET: 201 No. 5547859th Figueroa Street, Fifth Floor  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012-2628  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/101,593  
; FILING DATE: 19930802  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-305  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2760 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2760  
US-08-101-593-3

Query Match 3.2%; Score 37.2; DB 1; Length 2760;  
Best Local Similarity 44.9%; Pred. No. 0.35;  
Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagagaagaatttctgatactgataa 658  
DB 1376 ATGAAATATATCGACGAAACAGCTCTTAACCGAGTATGTAATCTGTGTTCCGAATG 1435  
QY 659 ttccgaattgaagatgaagacacgcagccatttgagcgaagaaggatgggaactt 718  
DB 1436 GATGAGTATGATTAACATCAAGAAAGTATCTTCCAAAGAAATCGCTAAAGTATTTT 1495  
QY 719 tgaattgggaagatltacgaagaatgaattcctcaattgtgattcttgaaggcactac 778

Db 1496 TCCAGCTGAAGACTGAAAAAATGTTCCGTGAAGAAATGATCCGAGCTATTA 1555  
Oy 779 gatggtgataatcgtcaagaactgacatagaagaactactcatatataatcaag 838  
Db 1556 AAAAATTTATTTATGAAGCGGAGGCTCTGTTCACAACTAAACCAAGTTGAACATATG 1615  
Oy 839 aatatgattccaaagggtggaagtggttcacatctccacagcagttacatctgac 898  
Db 1616 TTAAGTCAAGTATGATTTCTTAATGAACATTCGATTCACCGAATCTGTTCACATA 1675  
Oy 899 cctctctcatgaa 912  
Db 1676 GTCTGATTGAAGA 1689

RESULT 14  
US-08-323-170B-1  
Sequence 1, Application US/08323170B

Patent No. 5733772  
GENERAL INFORMATION:  
APPLICANT: Williamson, Kim C.  
APPLICANT: Kaslow, David C.  
TITLE OF INVENTION: Cloning and Expression of Plasmodium  
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf5230  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/323,170B  
FILING DATE: 13-OCT-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/010,409  
FILING DATE: 29-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Quine, Jonathan A.  
REGISTRATION NUMBER: P-41,261  
REFERENCE/DOCKET NUMBER: 015280-113100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0300  
TELEFAX: (415) 576-0200  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9636 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 149..9556  
US-08-323-170B-1

Query Match 3.2%; Score 37; DB 1; Length 9636;  
Best Local Similarity 52.2%; Pred. No. 0.69; 75; Indels 0; Gaps 0;  
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Oy 637 gaagatttcgattcgataattcgatgaagaatgaagaacatgacgacattagacc 696  
Db 4982 GAATCTATTATTCGAATCTTTAAGATTAATGAATCTAATGAATCTTTCACACC 5041  
Oy 697 aagaagggagtggaactttggaattgggaagatatcagaagatggaattactcaa 756

Db 5042 GGGGATGTAGATCGAAGAACATTTATATTAAGAGATATGATACATGGAAAGTTATAT 5101  
Oy 757 tgtgtgattctgaggaactagatggtgtaatcgc 793  
Db 5102 GATGAACATTAAGAAAGATTTATATATATGATATG 5138

RESULT 15  
US-08-954-441-1  
Sequence 1, Application US/08954441  
Patent No. 6316000

GENERAL INFORMATION:  
APPLICANT: Williamson, Kim C.  
APPLICANT: Kaslow, David C.  
TITLE OF INVENTION: Cloning and Expression of Plasmodium  
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf5230  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,441  
FILING DATE: 20-OCT-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/323,170  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/010,409  
FILING DATE: 29-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Einhorn, Gregory P.  
REGISTRATION NUMBER: 38,440  
REFERENCE/DOCKET NUMBER: 015280-113110US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0300  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9636 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 149..9556  
US-08-954-441-1

Query Match 3.2%; Score 37; DB 4; Length 9636;  
Best Local Similarity 52.2%; Pred. No. 0.69; 75; Indels 0; Gaps 0;  
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Oy 637 gaagatttcgattcgataattcgatgaagaatgaagaacatgacgacattagacc 696  
Db 4982 GAATCTATTATTCGAATCTTTAAGATTAATGAATCTAATGAATCTTTCACACC 5041  
Oy 697 aagaagggagtggaactttggaattgggaagatatcagaagatggaattactcaa 756  
Db 5042 GGGGATGTAGATCGAAGAACATTTAATTAATGAAGATGATGATCAGGAAAGTTATAT 5101  
Oy 757 tgtgtgattctgaggaactagatggtgtaatcgc 793

Tue Jul 30 08:40:09 2002

us-09-995-917a-3.rni

Page 9

Db 5102 GATGAACAATTAGAGAGAGTTATATATATATGATATG 5138

Search completed: July 29, 2002, 22:40:16  
Job time: 4302 sec

**THIS PAGE BLANK (USPTO)**